Impact of global changes on soil fungal diversity : an environmental genomics study

Claudia Bragalini

Université Claude Bernard Lyon1
Ecole Doctorale: Evolution, Ecosystèmes, Microbiologie et Modélisation

Università degli Studi di Torino
Scuola di Dottorato in Scienzedella Natura e Tecnologie Innovative

N° d'ordre: **40 - 2015**

Doctorat en Biologie Spécialité Microorganismes, Interactions, Infections (Arrêté du 7 Août 2006)
Discipline: "Ecologie Microbienne"

Tesi di Dottorato di Ricerca in Scienze Biologiche e Biotecnologie Applicate

Impact of global changes on soil fungal diversity. An environmental genomics study.

Impact des changements globaux sur la diversité des champignons du sol. Approche en génomique environnementale.

Impatto dei cambiamenti globali sulla diversità fungina del suolo. Uno studio di genomica ambientale.

1st April 2015

Claudia Bragalini

Members of the Thesis Board:

Rapporteurs:	Dr. Roberto GEREMIA	Research Director, CNRS, Grenoble, France
	Prof. Marco BAZZICALUPO	Full Professor, Univeristy of Florence, Italy
Examinateurs:	Prof. Mariangela GIRLANDA - Supervisor	Professor, University of Turin, Italy
	Dr. Roland MARMEISSE - Supervisor	Research Director, CNRS, Lyon, France
	Dr. Alberto ACQUADRO	Assistant Professors, University of Turin, Italy
	Dr. Duccio CAVALIERI	Research Director, Edmund Mach Foundation, Italy
	Prof. Ivan MOËNNE-LOCCOZ	Full Professor, Université C. Bernard Lyon1, France

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	**M. François-Noël GILLY**
Vice-président du Conseil d'Administration	M. le Professeur Hamda BEN HADID
Vice-président du Conseil des Etudes et de la Vie Universitaire	M. le Professeur Philippe LALLE
Vice-président du Conseil Scientifique	M. le Professeur Germain GILLET
Directeur Général des Services	M. Alain HELLEU

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur: M. le Professeur J. ETIENNE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur: Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur: M. le Professeur D. BOURGEOIS
Institut des Sciences Pharmaceutiques et Biologiques	Directeur: Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur: M. le Professeur Y. MATILLON
Département de formation et Centre de Recherche en Biologie Humaine	Directeur: Mme. la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur: M. F. DE MARCHI
Département Biologie	Directeur: M. le Professeur F. FLEURY
Département Chimie Biochimie	Directeur: Mme Caroline FELIX
Département GEP	Directeur: M. Hassan HAMMOURI
Département Informatique	Directeur: M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur: M. le Professeur G. TOMANOV
Département Mécanique	Directeur: M. le Professeur H. BEN HADID
Département Physique	Directeur: M. Jean-Claude PLENET
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur: M. Y.VANPOULLE
Observatoire des Sciences de l'Univers de Lyon	Directeur: M. B. GUIDERDONI
Polytech Lyon	Directeur : M. P. FOURNIER
Ecole Supérieure de Chimie Physique Electronique	Directeur: M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur: M. le Professeur C. VITON
Ecole Supérieure du Professorat et de l'Education	Directeur: M. le Professeur A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Directeur: M. N. LEBOISNE

Index

Abstract

Impact of global changes on soil fungal diversity. An environmental genomics study.

The soil environment is the most complex natural matrix, inhabited by extremely diverse organisms (belonging to almost all major prokaryotic and eukaryotic phyla), that contribute to a wide range of ecosystem services essential to the sustainable function of natural and managed ecosystems. It has been estimated that 1g of soil contains several billion bacteria with more than 10^3 different species, up to 50 km of fungal hyphae and several hundred species of fauna (protozoa, nematodes, insects, earthworms) also living in a small volume of soil (few cm^3). Fungi are a major component of soil biota. They exhibit huge species diversity and diverse life styles, playing fundamental ecological roles as either decomposers, mutualists or pathogens of plants and animals. They contribute to the maintenance the soil structure, drive carbon cycling in forest soil, mediate mineral nutrition of plants, and are involved in the regulation of the population and community structure of their plant or animal hosts. In addition to their ecological importance, due to their capacity to produce a wide spectrum of enzymes and small molecules, Fungi attract a wide biotechnological interest and are important in several industrial processes.

Global scenarios of biodiversity change in terrestrial biomes for the year 2100 have identified modification in land use as the driver that is expected to have the larger global impact on biodiversity, followed by climate change (e.g, elevated temperature or altered precipitation regimes), nitrogen deposition, invasion by exotic species and changes in atmospheric CO_2 concentration. Such factors have a potential impact on soil fungi. Understanding the factors driving soil biodiversity is of primary importance not only to predict the response of soil communities to such changes but also the impact that the latter will have on important ecosystems services. In this context, the general aim of this thesis was to assess the impact of two main recognized drivers of global changes on soil fungal communities.

One chapter (*Chapter II*) is dedicated to the impact of different land uses on the important group of obligate plant symbionts, the Arbuscular Mycorrhizal Fungi (AMF, *Glomeromycota)*. In the framework of the European project EcoFINDERS, the taxonomic diversity of AMF communities was investigated in soils from four different sites (Long Term Observatories, LTOs) across Europe. These LTOs represent habitats with different types of soil, land use and land use intensification level (high- or low-intensity) at both continental (among-LTO) and local (intra-LTO) scales. Soil samples were collected in spring and autumn 2011, and AMF communities were described by metabarcoding the ITS2 ribosomal gene region amplified from soil-extracted DNA followed by high-throughput (Roche 454 FLX) sequencing. At the continental scale, the land use intensification level *per se* and season did not significantly affect AMF community composition, while land use and LTO did. At the continental as well as at the local scale (for three out of the four LTOs) soil properties were identified as the factors which explained the highest amount of variance in AMF assemblages, by contrast in one LTO it was the season. A number of taxa/combinations of taxa were detected as indicators of land use/

combinations of land uses, but no LTO-independent indicators for either high- or low intensity was identified. These findings indicate that the effects of land use intensification on both individual AMF taxa and the entire communities are context-dependent. Adaptation to the local biotic environment and stochastic processes may also play important roles in shaping the communities of these symbiotic fungi.

The effect of climate changes, in particular of the reduction in precipitation predicted for the Mediterranean area, on soil fungi is the subject of the third chapter (*Chapter III*). The study took place at the Puéchabon experimental *Quercus ilex* forest (France), where a replicated experiment of reduction of throughfall precipitation has been conducted for several years. Soils samples were collected in two seasons (spring and autumn) between 2010 and 2012 (four sampling campaigns), in plots subjected to a reduction in throughfall precipitation (-29% average precipitation input to the soil), as well as in control plots. For the assessment of taxonomic and functional fungal community diversity, a parallel high-throughput (Illumina MiSeq) metabarcoding on soil-extracted RNA of four functional genes was performed (targeted-metatranscriptomics). One gene was representative of the global taxonomic fungal diversity (EF1-alpha), while the remaining three genes were involved in plant biomass degradation, a key process largely controlled by soil fungi, as they encode enzymes active on the three most abundant plant cell wall polymers (cellulose, hemicellulose and lignin). To target these gene families the design and optimization of gene-specific primers was necessary (*Chapter V*). The results obtained suggest that at Puéchabon, soil fungal communities are fairly stable components of the ecosystem, resistant to both temporal and/or environmental changes in term of alpha diversity. By contrast, sampling time (and to a lesser extent season, but not rainfall reduction) had a strong impact on beta-diversity indices for all the four studied genes. Two hypotheses may explain these observations: (i) the rainfall reduction treatment is in itself not perceived by soil microbial communities, or (ii) the microbial communities present in environments which naturally experience strong and recurrent climatic variations have developed adaptive strategies to cope with these variations and may be resistant, to some extent, to further increases in the latter. This is particularly relevant for water availability in Mediterranean areas which are naturally characterized by strong seasonal variations in the levels of precipitations.

Soil fungal communities potentially represent a rich source of novel natural products, including biocatalysts. Commonly used culture-independent PCR-based methods, as implemented in *Chapters II* and *III*, are powerful tools to describe functional gene diversity from uncultured microorganisms but they suffer from several limitations. Besides primers bias, generally they do not allow the recovery of full length genes. This is particularly true for functional genes for which primers are often designed on the more conserved catalytic domain sequence of the protein. Obtaining full-length functional genes from the environment is of particularly interest in the biotechnological field. In *Chapter IV* an original adaptation of the solution hybrid selection (SHS) for an efficient recovery of functional complementary DNAs (cDNAs) synthesized from soil-extracted polyadenylated mRNAs is described. This protocol was tested on the Glycoside Hydrolase 11 (GH11) gene family encoding endo-xylanases (involved in hemicellulose degradation and used in various

industrial applications). After two successive rounds of capture, almost all selected cDNAs (>90%) were GH11 sequences, of which 70% were full length. Moreover, a number (1.5%-25%) of the cloned captured sequences were successfully expressed in *Saccharomyces cerevisiae*. The sequencing of PCR-amplified GH11 gene fragments from the captured sequences highlighted hundreds of phylogenetically diverse sequences that were not yet described in public databases. These findings indicate that the protocol presented in *Chapter IV* offers the possibility of performing exhaustive exploration of truly functional fungal gene families in the environment.

Résumé

Impact des changements globaux sur la diversité des champignons du sol. Approche en génomique environnementale.

L'environnement sol représente la matrice environnementale la plus complexe. Elle est peuplée d'organismes très divers, appartenant à presque tous les grands phyla bactériens et eucaryotes, qui contribuent à un grand nombre de services écosystémiques essentiels pour un fonctionnement durable des écosystèmes naturels et anthropisés. Il a été estimé qu'un seul gramme de sol peut contenir plusieurs milliards de cellules bactériennes appartenant à plus de mille espèces, ainsi que jusqu'à 50 km d'hyphes fongiques. Plusieurs centaines d'espèces de protistes et d'animaux (nématodes, insectes, vers de terre) vivent aussi dans quelques cm^3 de sol. Les champignons représentent une composante essentielle de cette diversité. Ils montrent une grande diversité taxonomique à laquelle correspond une importante diversité fonctionnelle et de modes de vie qui conduisent ces organismes à jouer des rôles fondamentaux comme décomposeurs, symbiontes ou pathogènes de plantes et d'animaux. Ils contribuent au maintien de la structure du sol, ils contrôlent le recyclage du carbone, contribuent à la nutrition minérale des plantes et interviennent dans la régulation des populations et des communautés de leurs hôtes végétaux ou animaux. Au-delà de cette importance écologique, du fait de leur capacité à produire une large gamme d'enzymes et de micromolécules, les champignons représentent un intérêt en biotechnologie et sont importants pour la réalisation de différents procédés industriels.

Les scénarios globaux sur les changements de biodiversité dans les biomes terrestres à l'horizon 2100 identifient les modifications dans l'usage des terres comme le facteur qui aura le plus important impact sur la biodiversité, suivi des changements climatiques (parmi lesquels l'élévation des températures ou l'altération des régimes de précipitations), de la déposition d'azote, des disséminations d'espèces et des changements du taux de CO_2 atmosphérique. Ces facteurs ont un impact potentiel sur les champignons du sol. La compréhension des facteurs contrôlant la biodiversité des sols est donc de première importance non seulement pour prédire la réponse des communautés microbiennes des sols à ces changements, mais aussi pour apprécier l'impact que ces derniers auront sur les services écosystémiques. Dans ce contexte, le but général de cette thèse a été d'évaluer l'impact de deux facteurs des changements globaux sur les communautés de champignons du sol.

Un chapitre (*Chapitre II*) est consacré à l'impact de différents usages des terres sur un groupe important de symbiontes obligatoires des plantes, les champignons mycorhiziens à arbuscules (champignons AMF, du groupe des Glomeromycota). Dans le cadre du projet Européen EcoFINDERS, la diversité taxonomique des communautés d'AMF a été étudiée dans des sols de 4 sites d'observation à long terme (LTO) en Europe. Ces LTO englobent des habitats caractérisés par différents types de sols, d'usages des terres ainsi que "d'intensités" dans leur usage (intensités élevées et faibles) à des échelles continentales (entre LTO) et locales (au sein même des LTO). Les sols ont été collectés au printemps et à l'automne 2011 et les communautés d'AMF décrite par "métabarcoding" de la région

ITS2 de l'ARN ribosomique amplifiée de l'ADN extrait directement du sol et séquencée par "pyroséquençage 454". A l'échelle continentale, l'intensification de l'usage des terres *per se* et la saison n'affectent pas significativement la composition des communautés d'AMF, tandis que l'usage des terres et le LTO ont un effet significatif. Aux échelles continentales et locales (pour 3 des 4 LTO), les caractéristiques des sols ont été identifiées comme étant les facteurs expliquant la plus forte proportion de la variation totale au sein des assemblages d'espèces AMF. Pour un des LTO le facteur explicatif majeur était toutefois la saison. Un certain nombre de taxa ou de combinaisons de taxa ont été identifiés comme indicateurs d'un usage particulier ou de plusieurs usages des terres; mais aucun indicateur indépendant des LTO n'a pu être identifié pour, soit un fort niveau, soit un faible niveau d'intensité d'usage des terres. Ces résultats indiquent que les effets de l'intensification de l'usage des terres sur à la fois des taxa particuliers d'AMF ou sur leurs communautés dans l'ensemble sont dépendant du contexte local. L'adaptation à l'environnement biotique local et des processus stochastiques pourraient aussi jouer des rôles importants dans l'assemblage des communautés de champignons symbiotiques.

L'effet des changements climatiques, en particulier la réduction des précipitations prédite pour le bassin Méditerranéen, sur les champignons du sol est le sujet du 3ème chapitre. Cette étude a été réalisée sur le site expérimental de la forêt de chênes verts de Puéchabon (France) où une expérimentation répliquée de réduction des précipitations incidentes est conduite depuis plusieurs années. Des échantillons de sols ont été prélevés sur 2 saisons (printemps et automne) entre 2010 et 2012 (4 campagnes d'échantillonnage), dans des parcelles sujettes à une réduction de la pluviométrie (-29%) et dans des parcelles témoins. Pour l'estimation de la diversité taxonomique et fonctionnelle des communautés fongiques, le "métabarcoding" à haut débit de 4 gènes a été réalisé à partir des ARN extraits de sol (métatranscriptomique). Un gène (EF1α) est représentatif de la diversité taxonomique globale des champignons. Les 3 autres gènes sont impliqués dans la dégradation de la biomasse végétale, un processus clé largement contrôlé par les champignons, et codent des enzymes actives sur les 3 polymères les plus abondants de la paroi végétale (la cellulose, l'hémicellulose et la lignine). Le dessin et l'optimisation d'amorces PCR spécifiques de ces gènes a été préalablement nécessaire (*Chapitre V*). Les résultats obtenus montrent que les niveaux de diversité alpha des communautés fongiques sont relativement stables et ne varient pas significativement au cours du temps et en réponse aux variations de précipitation. Toutefois, la date de prélèvement (et dans une moindre mesure la saison, mais pas la réduction de pluviométrie) a un fort impact sur les indices de diversité béta pour les 4 familles de gènes étudiées. Deux hypothèses peuvent être formulées pour expliquer ces observations. (i) La réduction de pluviométrie appliquée sur les parcelles n'est pas suffisamment intense pour être perçue par les communautés microbiennes du sol ou (ii) les communautés microbiennes présentes dans des environnements naturellement caractérisés par des variations climatiques importantes et récurrentes ont développées des stratégies adaptatives pour répondre à ces variations et pourraient être résistantes, jusqu'à un certain point, à

une augmentation de l'intensité des variations climatiques. Ceci est applicable à la disponibilité en eau dans les zones méditerranéennes qui sont naturellement caractérisées par de fortes variations saisonnières des niveaux de précipitations.

Les communautés fongiques des sols représentent une source potentiellement riche en nouveaux produits naturels, dont des biocatalyseurs. Les méthodes de PCR, indépendantes des mises en cultures (comme mises en œuvre dans les *Chapitres II et III*) sont des outils puissants permettant de décrire la diversité fonctionnelle des microorganismes non cultivables. Mais ces approches ont des limites, notamment elles ne permettant pas l'obtention de gènes entiers dans la mesure ou les amorces de PCR sont généralement conçues sur la base des domaines catalytiques internes aux protéines. L'obtention de séquences entières est pourtant un enjeu en termes d'applications biotechnologiques de la biodiversité. Dans le *Chapitre IV*, est décrite une adaptation originale du protocole de capture de molécules hybrides en solution pour un isolement efficace d'ADNc entiers à partir d'ARNm polyadénylés directement extraits de sols. Ce protocole a été mis en œuvre sur la famille 11 de glycoside hydrolases (GH11) codant des endoxylanases, impliquées dans la dégradation de l'hémicellulose et utilisées dans différents procédés industriels. Après deux étapes successives de capture, presque tous les ADNc sélectionnés (>90%) étaient des séquences GH11, parmi lesquelles 70% étaient entières. De plus, un nombre significatif (entre 1,5 et 25%) des séquences capturées puis clonées ont été exprimées avec succès dans la levure *Saccharomyces cerevisiae*. Le séquençage des fragments GH11 amplifiés des séquences capturées identifie des centaines de séquences phylogénétiquement diverses non présentes dans les bases de données publiques. Ces résultats démontrent que le protocole décrit dans le *Chapitre IV* offre la possibilité de réaliser une exploration exhaustive de familles de gènes fongiques exprimés dans l'environnement.

Riassunto

Impatto dei cambiamenti globali sulla diversità fungina del suolo. Uno studio di genomica ambientale.

In virtù della sua eterogeneità, il suolo è l'ambiente naturale più complesso, colonizzato da organismi estremamente diversi (appartenenti a quasi tutti i principali phyla procariotici ed eucariotici), i quali contribuiscono ad un ampio spettro di servizi essenziali per il funzionamento di ecosistemi naturali ed agroecosistemi. E' stato stimato che un grammo di suolo può contenere miliardi di batteri (oltre 10^3 specie) e fino a 50 km d'ife fungine. Un piccolo volume di suolo (pochi cm^3) ospita inoltre diverse centinaia di specie di protozoi e animali (nematodi, insetti, lombrichi). I Funghi rappresentano una componente importante degli organismi tellurici. Questi microrganismi sono caratterizzati da un'enorme diversità di specie e stili di vita, e svolgono ruoli ecologici fondamentali in qualità di decompositori, simbionti mutualisti o patosisti di piante e animali. Essi contribuiscono al mantenimento della struttura del suolo, guidano il ciclo del carbonio nei suoli forestali, mediano la nutrizione minerale delle piante, intervengono nella nutrizione di altri organismi del suolo e sono coinvolti nella regolazione della struttura di popolazioni e comunità dei loro ospiti vegetali e animali. Oltre alla loro importanza ecologica, grazie alla capacità di produrre un ampio spettro di enzimi e piccole molecole, i funghi rivestono interesse biotecnologico e sono importanti in diversi processi industriali.

Le proiezioni globali per il 2100 hanno identificato le variazioni d'uso del territorio come il fattore che è atteso avere il maggiore impatto sulla biodiversità terrestre, seguito dai cambiamenti climatici (quali variazioni termiche o alterazioni nelle precipitazioni), dall'incremento dei tenori in azoto del suolo, dall'invasione di specie esotiche e dalle variazioni nella concentrazione atmosferica di CO_2. Tutti questi fattori hanno un potenziale impatto sui funghi del suolo. Una migliore comprensione dei processi che regolano la biodiversità del suolo è di primaria importanza per predire gli effetti dei cambiamenti globali non solo sulle comunità di organismi tellurici, ma anche su importanti servizi ecosistemici. In questo contesto, scopo generale di questa tesi è stato quello di studiare l'impatto di due fattori, riconosciuti tra i maggiori responsabili dei cambiamenti globali, sulle comunità fungine del suolo.

Un capitolo della tesi (*Capitolo II*) è dedicato allo studio degli effetti di differenti usi del territorio su un importante gruppo di simbionti obbligati delle piante, i funghi micorrizici arbuscolari (Arbuscular Mycorrhizal Fungi, AMF; Glomeromycota). Nell'ambito del progetto europeo EcoFINDERS, è stata indagata la diversità sistematica delle comunità di AMF in suoli provenienti da quattro siti europei (Long Term Observatories, LTOs), rappresentativi di habitats caratterizzati da differenze, su scala sia continentale (tra LTOs) sia locale (intra-LTO), nel tipo di suolo, nell'uso del territorio, e nei livelli d'intensificazione. I campioni di suolo sono stati raccolti in primavera e autunno 2011, e le comunità AMF sono state descritte tramite metabarcoding della regione genica ribosomale ITS2, amplificata da DNA estratto dal suolo, e successivo sequenziamento "massivo" (Roche 454 FLX). Su scala

continentale, il livello d'intensificazione e le variazioni stagionali non hanno determinato effetti significativi sulla composizione delle comunità AMF, mentre tali effetti sono stati osservati per le diverse tipologie di uso del territorio ed i diversi LTO. Su scala sia continentale sia locale (per tre dei quattro LTOs), i parametri fisico-chimici del suolo sono risultati essere i fattori in grado di spiegare la maggior parte della varianza nella struttura delle comunità AMF (per il restante LTO, al contrario, il fattore principale è risultato essere la stagione). Alcuni taxa/combinazioni di taxa sono stati identificati come indicatori di specifici usi/combinazioni di usi del territorio, ma non è stato identificato nessun indicatore LTO-indipendente (nè per alti nè per bassi livelli d'intensificazione). Questi risultati suggeriscono che gli effetti dell'intensificazione siano dipendenti dal contesto (sia per singoli taxa AMF sia per le comunità in toto). L'adattamento all'ambiente biotico locale e processi stocastici contribuiscono, con ogni probabilità, a strutturare le comunità di questi funghi simbionti nel suolo.

Oggetto del terzo capitolo (*Capitolo III*) è l'effetto dei cambiamenti climatici, in particolare della riduzione delle precipitazioni predetta per l'area mediterranea, sui funghi del suolo. Lo studio è stato condotto nella lecceta sperimentale di Puéchabon (Francia), dove un esperimento di riduzione delle precipitazioni è stato allestito per diversi anni. I campioni di suoli analizzati sono stati raccolti in due stagioni (primavera e autunno), tra il 2010 e il 2012 (quattro campagne di campionamento), sia da plots sottoposti a un trattamento di riduzione delle precipitazioni (riduzione media del 29% delle precipitazioni al suolo) sia da plots di controllo. Lo studio della diversità sistematica e funzionale delle comunità fungine è stato condotto attraverso un approccio di metabarcoding su RNA estratto da suolo (metratrascrittomica mirata), seguito da sequenziamento "massivo" (Illumina MiSeq) di quattro geni funzionali. Il gene EF1-alpha è stato scelto per rappresentare la diversità fungina globale; i rimanenti tre geni sono invece coinvolti nella decomposizione della biomassa vegetale (processo largamente controllato dai funghi nel suolo), dal momento che codificano enzimi attivi su tre dei principali polimeri della parete vegetale (cellulose, emicellulose e lignina). Per analizzare queste famiglie geniche è stato necessario disegnare e validare primers specifici (*Capitolo V*). I risultati ottenuti suggeriscono che nel suolo della foresta di Puéchabon le comunità fungine sono componenti stabili dell'ecosistema, resistenti a cambiamenti temporali e/o ambientali, in termini di alfa-diversità. I cambiamenti temporali (giorno di campionamento ed, in misura minore, stagione) hanno invece determinato effetti significativi sugli indici di beta-diversità relativi a tutti i geni target. Queste osservazioni potrebbero essere spiegate da due ipotesi: (i) il trattamento di riduzione delle precipitazioni non è percepito dalle comunità fungine, oppure (ii) le comunità fungine di ambienti soggetti a variazioni climatiche naturali forti e ricorrenti hanno sviluppato strategie adattative per far fronte a tali variazioni e possono quindi risultare resistenti ad ulteriori cambiamenti. Quest'ultimo scenario sembra particolarmente plausibile per quanto attiene alla disponibilità idrica nelle aree mediterranee, che sono naturalmente caratterizzate da forti fluttuazioni stagionali nelle precipitazioni.

Le comunità fungine del suolo rappresentano una ricca fonte potenziale di nuovi prodotti naturali, inclusi biocatalizzatori. Gli approcci sperimentali

convenzionalmente adottati per descrivere la diversità funzionale dei microorganismi non coltivabili (quali quelli usati per le indagini descritte nel *Capitolo III*), sono basate sulla PCR, e soffrono pertanto di alcune limitazioni. Oltre a risentire del "primers bias", i metodi PCR-dipendenti non permettono generalmente di ricostruire l'intera sequenza genica. Questa limitazione si applica soprattutto ai geni funzionali, i primers per amplificare i quali sono spesso disegnati sulle sequenze conservate presenti nel dominio catalitico della proteina. A fini biotecnologici è tuttavia d'interesse ottenere geni interi dai campioni ambientali. Nel *Capitolo IV* è descritto un adattamento originale della tecnica "solution hybrid selection" (SHS), sviluppato allo scopo di ottenere una resa efficiente di DNA complementari (cDNAs) funzionali sintetizzati a partire da mRNA poliadenilato estratto da suolo. Il protocollo descritto è stato sviluppato sulla famiglia genica *Glycoside Hydrolase* 11 (GH11), codificante per endo-xylanasi (enzimi coinvolti nella degradazione delle emicellulose e usati in varie applicazioni industriali). Dopo due cicli successivi di "capture", la quasi totalità dei cDNAs ottenuti (>90%) sono risultati corrispondere a sequenze GH11, ed il 70% di tali sequenze geniche erano intere. Inoltre, una buona proporzione (1.5%-25%) delle sequenze recuperate con la "capture" e clonate sono state espresse con successo in *Saccharomyces cerevisiae*. Il sequenziamento di frammenti GH11 amplificati tramite PCR, recuperati con la "capture", ha rivelato l'esistenza di centinaia di sequenze filogeneticamente distinte da quelle presenti nelle banche dati pubbliche. I risultati ottenuti indicano come il protocollo descritto nel *Capitolo IV* offra la possibilità di esplorare in modo esaustivo la diversità naturale di famiglie geniche fungine realmente funzionali.

Chapter I

General introduction

1.1 The dimension of fungal occurrence and diversity in soil

Because of its extreme physical and chemical heterogeneity at small scales and microclimatic characteristics, the soil environment is the most complex natural matrix, inhabited by extremely diverse organisms (Jeffery *et al.* 2010). It has been estimated that 1g of soil contains several billion bacteria with more than 10^3 different species, up to 50 km of fungal hyphae and several hundred species of fauna (protozoa, nematodes, insects, earthworms) also living in a small volume of soil (few cm^3) (Bardgett 2005, Roesch *et al.* 2007, Jeffery *et al.* 2010). Encompassing an estimated 55-98% of total biodiversity on Earth (Beed *et al.* 2011), few other ecosystems can therefore compare in size, complexity or diversity with soil.

Fungi are a major component of the soil biota. The mycelial growth-form is well adapted to the soil environment, since hyphae can effectively explore the three-dimensional soil pore network, foraging for food resources. A number of biomarkers and methods have been used to quantify production, standing biomass and turnover of mycelia of either symbiotic and saprotrophic fungi in the field (e.g. Leake *et al.* 2004a, Strickland & Rousk 2010, Ekblad *et al.* 2013, Wallander *et al.* 2013). Although estimates depend on the applied methods and techniques (Wallander *et al.* 2013), in some soils (such as forest soils) mycelia have been found to account for most microbial biomass, outcompeting bacteria (Strickland & Rousk 2010). For instance, the production rate of (mycorrhizal) mycelium in the upper 10 cm of ~140 different forest sites was found to average 160 kg dry matter ha^{-1} $year^{-1}$ (Ekblad *et al.* 2013), reaching up to 980 kg dry matter ha^{-1} over 4 months in a *Pinus taeda* plantation at low elevation in North Carolina (Parrent & Vilgalys 2007). A comparison of the mineralization rates of chitin and peptidoglycan shows that chitin has a mineralization rate equal to or ~30% lower than peptidoglycan (Li & Brune 2005a, 2005b). Consequently, hyphae are generally more persistent in the soil than bacterial cells (Martin *et al.* 1979, Solomon *et al.* 2001, Amelung *et al.* 2002, but see De Vries *et al.* 2007). Aboveground plant litter quality and decomposition rates have been proposed as the fundamental determinants of long-term soil organic matter (SOM) accumulation (e.g. Wardle *et al.* 2003, Brovkin *et al.* 2012, Makkonen *et al.* 2012). However, using ^{14}C bomb-carbon modeling, a recent study (Clemmensen *et al.* 2013) showed that 50 to 70% of stored carbon in a chronosequence of boreal forested islands derives from roots and root-associated fungi, pointing to impaired decomposition and preservation of fungal residues as an important regulator of C accumulation in late successional forests.

Although fungal biomass is prominent and persistent in (at least certain) soils, the actual dimension of fungal diversity in this environment has only been recently investigated with culture-independent molecular methods taking advantage of high-throughput sequencing, which outperform earlier approaches in terms of resolution and magnitude (e.g. Buée *et al.* 2009, Tedersoo *et al.* 2010, Lentendu *et al.*

2011, Daghino *et al.* 2012, Orgiazzi *et al.* 2012, 2013; Schmidt *et al.* 2013, Pellissier *et al.* 2014, Taylor *et al.* 2014, Timling *et al.* 2014). Traditionally, most effort in describing and cataloguing fungal species has been directed towards the larger and more "visible" species (Bass & Richards 2011). However, environmental sequencing has revealed far higher fungal species-level diversity than suggested by their morphological diversity. From the sequencing of cloned environmental sequences, O'Brien *et al.* (2005) estimated 491 fungal OTUs (Operational Taxonomic Units) in their pine forest soil samples and 616 in mixed hardwood plot samples. These analyses were performed on a few grams of soil and resulted in an underestimation of the local diversity as the ACE richness estimator continued to increase beyond the limits of the sampling effort. A pilot study making use of 454 pyrosequencing to evaluate the fungal diversity in six distinct and spatially distant soil samples from a temperate forest (Buée *et al.* 2009) recovered approximately 30,000 ITS reads in each forest soil sample (4 g), corresponding to about 1000 molecular OTUs. A concordant level of diversity was reported by Lentendu *et al.* (2011) for alpine soil samples. Taylor and his colleagues (2010) found more than 200 operational taxonomic units in a 0.25 g soil sample with only 14% overlap in a sample taken a meter away. Taylor *et al.* (2014) achieved the first exhaustive enumeration of fungi in soil, recording 1002 taxa in *Picea mariana* forest soils from interior Alaska. These studies reveal the existence of a fungal rare biosphere, and by retrieving a high proportion of novel fungal sequence types, they challenge our understanding of global fungal biodiversity. Until recently, indeed, estimates of numbers of fungi did not include results from environmental sequencing methods. In 1991, a landmark paper provided several qualified estimates of the number of fungi on the Earth based on ratios of known fungi to plant species in regions where fungi were considered to be well-studied (Hawksworth 1991). This estimate of 1.5 million species was accepted as a reasonable working hypothesis. However, Hawksworth's (1991) estimate is now considered to be conservative by many, including Hawksworth (Hawksworth 2012). More recent estimates based on data acquired from several molecular methods have predicted as many as 5.1 million species of fungi (O'Brien *et al.* 2005, Taylor *et al.* 2010, Blackwell 2011, Bass & Richards 2011, Hibbett & Taylor 2013, Money 2013). The actual dimension of the ''rare biosphere'' has been debated, based on the possible artifactual nature of "singletons", i.e. OTUs that include unique reads (Quince *et al.* 2009, Reeder & Knight 2009, Dickie 2010, Tedersoo *et al.* 2010). However, concomitant high-throughput sequencing and fungal isolation provides evidence that OTUs represented by a single read may correspond to real biological entities (Daghino *et al.* 2012).

Environmental studies identified not only novel individual species, but also major clades of fungi, such as the class Archaeorhizomycetes (Rosling *et al.* 2011), containing a diverse group of soil-inhabiting fungi from the phylum Ascomycota.

Sequences of Archaeorhizomycetes members have been reported in more than 50 independent studies, and they can be grouped into more than 100 species-level entities (Rosling *et al.* 2011). Nevertheless, only one species, *Archaeorhizomyces finlayi*, has been formally described, based on a culture that was obtained from conifer roots. A similar example is provided by the phylum Rozellomycota (James & Berbee 2012) (also known as Cryptomycota; Jones *et al.* 2011), a large clade of aquatic and soil-inhabiting fungi that is known almost exclusively from environmental sequences (Lara *et al.* 2010, Jones *et al.* 2011). The phylum Rozellomycota has been shown to contain the previously described chytrid genus *Rozella* (Jones *et al.* 2011), but most of the diversity of this phylum resides in groups that are known only from environmental sequences and have not been named. These new groups comprise physically small, cryptic, and elusive elements of fungal diversity. This diversity - in terms of rDNA variation - can be huge, as demonstrated for the Cryptomycota by Jones *et al.* (2011). The recognition of Cryptomycota alone could radically increase the size of the fungal kingdom. Furthermore, there is increasing evidence and consensus that other very diverse group of elusive organisms, the protist-like Microsporidia and Aphelidea, also belong to the Fungi (Keeling 2003; Lee *et al.* 2008, 2010; Karpov *et al.* 2013, 2014, James *et al.* 2013). These examples, and many others from fungal molecular ecology, illustrate the profound disconnection existing between formal taxonomy and species discovery through environmental sequencing (Hibbett & Taylor 2013), offering robust reasons to reconsider fungal diversity in soil.

1.2 The interactions of soil fungi with plants

Due to their high diversity, genetic plasticity and physiological versatility, soil fungi are the foundation of many ecosystems services ('the benefits people obtain from ecosystems', Millenium Ecosystem Assessment 2003; Table 1), that are used by human society for its own purposes, such as food production and climate regulation (Kibblewhite *et al.* 2008; Turbe´ *et al.* 2010; Dominati *et al.* 2010).

As indicated in Table 1, many ecosystem services in terrestrial environments are provided by fungi that positively interact with plants – either mycorrhizal fungi (establishing symbiotic relationships with living plants), or saprotrophic fungi (which decompose dead plant remains).

Table 1. Ecosystem services provided by fungi (modified from Dighton 2003)

Ecosystem service		Fungal functional group
Soil formation	Rock dissolution	Lichens, Saprotrophic fungi, Mycorrhizal fungi
	Particle binding	Saprotrophic fungi, Mycorrhizal fungi
Providing fertility for primary production	Decomposition of organic residues	Saprotrophic fungi, Mycorrhizal fungi
	Nutrient mineralization	Saprotrophic fungi, Mycorrhizal fungi
	Soil stability (aggregates)	Saprotrophic fungi, Mycorrhizal fungi
Primary production	Direct production	Lichens
	Nutrient accessibility	Mycorrhizal fungi
	Plant yield	Mycorrhizal fungi, Pathogenic fungi
	Defense against pathogens	Mycorrhizal fungi, Endophytic fungi, Pathogenic fungi
	Defense against herbivory	Endophytic fungi
Plant community structure	Plant-plant interactions	Mycorrhizal fungi, Pathogenic fungi
Secondary production	As a food source	Saprotrophic fungi, Mycorrhizal fungi
	Population/biomass regulation	Pathogenic fungi
Modification of pollutants		Saprotrophic fungi, Mycorrhizal fungi
Carbon sequestration and storage		Saprotrophic fungi, Mycorrhizal fungi

1.2.1 Mycorrhizal fungi

The term mycorrhiza was coined in the last part of the 19th century to design symbioses between plant roots and soil fungi (Smith & Read 2008). Mycorrhizal fungi are specialized root symbionts, engaging in intimate association with a great diversity and majority of land plants in all terrestrial ecosystems around the globe (Smith & Read 2008). Historically , the variety of mycorrhizal associations established between plants and fungi has been placed into seven categories, the main ones being the ectomycorrhiza (EM) and the arbuscular mycorrhiza (AM), based essentially on the structural characteristics of the symbiotic interfaces and the taxonomic identity of the symbionts (Smith & Read 2008). Specifically, about 80% of the extant land plants form AM symbiosis, with obligate symbiotic fungi which have been reclassified on the basis of DNA sequences into a separate fungal phylum, the Glomeromycota (Schüßler *et al.* 2001). It seems highly likely that the fungi had their origins possibly over 1000 million years ago (predating current estimates of

colonization of land by plants) and that AM symbioses are also extremely ancient (Redecker 2000). Through their roles in nutrient uptake, AM fungi were probably important in the colonization of land by plants; they remain major determinants of plant interactions in ecosystems to the present day (Smith & Read 2008). As opposed to the AM symbiosis, established exclusively by a specific fungal group, the ectomycorrhizal (ECM) one involves a diverse range of fungi belonging to the Ascomycota and Basidiomycota and it emerged long after the AM association (Bruns & Shefferson 2004). The ECM symbiosis concerns essentially ligneous plants, among which a majority of forest tree species in the boreal and temperate regions of the globe (Smith & Read 2008).

The best understood function of the fungi in the symbiosis is the improvement of plant mineral nutrient acquisition in exchange for up to 10-30% of the host plant photosynthates (Hobbie 2006), resulting in positive host growth responses. However, this symbiosis has a multifunctional character because mycorrhizal fungi may perform many other significant roles, including protection of the plant from biotic and abiotic stresses, for instance by altering host environmental tolerances to water deficit or pollutants, or reducing susceptibility to soil-borne pathogens. Although the association is generally assumed to be mutualistic, with bilateral nutrient exchange between plant and fungal partners, host responses ranging from positive to negative may in fact be observed, with mycorrhizal fungi sometimes functioning as commensals, necrotrophs or antagonists of host or non-hosts plants, their roles varying during the lifespan of the association (Johnson *et al.* 1997; Brundrett 2002, 2004; Egger & Hibbett 2004).

Recent work has provided clues on the selective forces maintaining cooperation between plants and mycorrhizal fungi. Plants and their mycorrhizal fungal symbionts interact in complex underground networks involving multiple partners. Each individual may indeed interact with several partners simultaneously: multiple mycorrhizal fungi can colonize a single plant host, and each of these fungal individuals can potentially interact with multiple plant hosts (Selosse *et al.* 2006; Kiers & Denison 2008). This increases the potential for exploitation and defection by individuals, raising the question of how partners maintain a fair, two-way transfer of resources (Kiers *et al.* 2011). By tracking the incorporation of carbon from *Medicago truncatula* plants into RNAs of fungal partners (three arbuscular mycorrhizal fungal species within the cosmopolitan subgenus *Glomus*), Kiers and colleagues (2011) observed that the most "cooperative" fungi (i.e. those transferring more nutrients) receive more carbon from root cells than the other species, showing that plants can detect, discriminate, and reward fungal partners at a fine scale, even when multiple fungi colonize a root. The authors also found that in turn, the fungal partners enforce cooperation by increasing phosphorus transfer only to those roots providing more carbohydrates. These observations indicate that the stability of the arbuscular

mutualism arises in a different way compared with other mutualisms, in which one partner appears to be "in control" and has either domesticated the other partner or enforces cooperation through punishment or sanction mechanisms. In these cases, the potential for enforcement has only been found in one direction, with the controlling partner housing the other partner in compartments, which can be preferentially rewarded or punished, such as in legume root nodules (legume species invest fewer resources into root nodules containing rhizobia that fix less nitrogen; Kiers *et al.* 2003, Simms *et al.* 2006, Oono *et al.* 2009, 2011, Regus *et al.* 2014). In contrast, in the mycorrhizal mutualism, both sides interact with multiple partners, so that neither partner can be "enslaved", and rather, the mutualism is evolutionarily stable because control is bidirectional, and both partners are able to preferentially reward the other. This provides an example of how cooperation can be stabilized in a form analogous to a market economy, where there are competitive partners on both sides of the interaction and higher quality services are remunerated in both directions, thus leading to the "biological market" metaphor to describe the mycorrhizal symbiosis (Kiers *et al.* 2011, Selosse & Rousset 2011, Wyatt *et al.* 2014).

Extraradical mycelia of mycorrhizal fungi are normally the "hidden half" of the symbiosis, but they exert powerful underground influences upon biogeochemical cycling, the composition of plant communities, and agroecosystem functioning (Leake *et al.* 2004a). The huge and extensive web of underground mycorrhizal hyphae in soils stores substantial amount of carbon biomass, besides constantly draining large amounts of carbon from their hosts again the exchange of other benefits. The development of techniques such as mycelium in-growth bags, chemical, molecular or isotopic markers, as well as large scale manipulations such as trenching and girdling experiments (Nylund & Wallander 1992, Ekblad & Näsholm 1996, Ekblad *et al.* 1998, Wallander *et al.* 2001, Dickie *et al.* 2002, Johnson *et al.* 2002, Leake *et al.* 2006, Högberg *et al.* 2010, Heinemeyer *et al.* 2012), which can distinguish between mycorrhizal extraradical mycelium and the mycelium of saprotrophic fungi, allowed to assess the extent directly *in-situ*. These techniques have revealed that the mycorrhizal mycelium is dominant in the soil of many terrestrial ecosystems, and mycorrhizal fungi have been identified as the main pathway by which recently fixed C enters soils (Godbold *et al.* 2006, Clemmensen *et al.* 2013, see also paragraphs 1.1).

A breakthrough in mycorrhizal ecology has been the discovery that individual mycelia of either ECM or AM fungi can interlink different host plants, thus establishing common mycelial networks (CMNs, often referred to as the 'wood-wide web' in the case of ECM mycelia) that connect plants, belonging to the same or different species, and providing potential pathways for interplant transport of mineral nutrients and C (Simard *et al.* 2002, Simard & Durall 2004, Taylor 2006, Selosse *et al.* 2006). Given the low specificity of many species of mycorrhizal fungi, shared symbionts between plants of different species might be common in nature. It

is then possible that carbon and nutrients might be transferred from plant to plant through the CMNs (Simard & Durall 2004, Selosse *et al.* 2006, van der Heijden & Horton 2009), and this could alter plant competitive ability. Evidence of a functional role of CMNs in nutrient transport and exchange derives from isotope labelling studies. Direct transfer of resources *via* both AM and ECM CMNs has been proposed following experiments showing that mineral resources such as N and P labelled in one plant can be detected in a second individual (Simard *et al.* 1997, 2002, Fitter *et al.* 1998, He *et al.* 2003, 2005, 2006, Simard & Durall 2004, Wilson *et al.* 2006). This indicates that mineral resources such as N and, to a lesser extent, P move between plants via mycorrhizal networks (Simard *et al.* 2002, Tuffen *et al.* 2002, He *et al.* 2003, 2005, 2006). Carbon also moves within both ECM and AM CMNs, but the actual net transport of C between interconnected plants remains controversial (Robinson & Fitter 1999, Simard *et al.* 2002, Simard & Durall 2004, Taylor 2006, Selosse *et al.* 2006). Transfer of resources via CMNs, if supported, reflects a unique feature of the mycorrhizal symbiosis. Should the resources flow from plants with higher to plants with lower levels of resources (sometimes referred to as 'source–sink'), resource transfer could potentially contribute to plant species coexistence through minimization of differential access to resources, with linked plant species forming guilds of mutual aid (Simard & Durall 2004, van der Heijden & Horton 2009). Therefore, the possibility of substantial interplant C transfer has led to the hypothesis that such transfers may influence interactions in plant communities, thus suggesting the need for a radical reappraisal of conventional concepts of competition in plant ecology (Leake *et al.* 2004a). Whereas there is clear evidence of net movement of carbon to mycoheterotrophic, non-photosynthetic plants which parasitize fungi for carbon (Leake 1994, Taylor *et al.* 2002, Leake 2004b, Bidartondo 2005), there is little evidence of ecologically meaningful exchange of resources between photosynthetic plants or that there is a significant net directional flow, as predicted by a source–sink relationship (Robinson & Fitter 1999, Wilson *et al.* 2006, Nakano-Hylander & Olsson 2007). Where quantification of the extent of resource transfer has been achieved, quantities transferred can be very low, representing as little as 0.004% of photosynthetic carbon gain (Teste *et al.* 2010). Furthermore, defoliation of adult plants has been found to contribute to an increased mycorrhizal benefit in neighboring seedlings, the opposite of the prediction of source–sink relationships (Pietikainen & Kytoviita 2007). In another study (Teste *et al.* 2010), larger planted seedlings have been found to receive more carbon transfer than smaller seedlings, and the amount of carbon fixed by donor plants has been found to be unrelated to transfer Finally, C provided by a donor plant may move to the roots of a recipient plant via a CMN, but remain in fungal compounds, indicating that carbon taken up by the mycorrhizal fungus in association with one mycorrhizal root does not become nutritionally available to other roots (Pfeffer *et al.* 2004).Before

resource transfer via CMNs can be incorporated into more general theories of plant coexistence, the ecological significance of any potential shared C resources between plants must be quantitatively demonstrated to cause an increase in plant performance (Bever *et al.* 2010).

Another way by which modification of resource access mediated by mycorrhizal fungi can shape plant species coexistence is by altering resource partitioning among plants. Prevailing models explaining species coexistence predict that competing species can coexist provided they are most limited by different resources and that they consume the resource they are most limited by at a higher rate than do other species. Since mycorrhizal symbionts modify nutrient uptake, they should also modify the conditions for competitive coexistence, either positively or negatively. A mechanism through which fungal-mediated resource partitioning could contribute to plant species coexistence involves plant species associating with different fungal symbionts which then provide differential access to alternate forms of particular resources. Individual species of ectomycorrhizal fungi, for example, can preferentially associate with specific hosts (e.g. Tedersoo *et al.* 2008) and vary in their access to mineral and organic forms of N and P (e.g. Tibbett & Sanders 2002). It is then possible that preferential association within this symbiosis directly contributes to resource partitioning of their hosts.

In the frameworks discussed above (mycorrhizal fungal mediation of soil resource partitioning and the CMNs), plant community dynamics are driven by resource competition or sharing of resources, respectively. The composition of the fungal community is critical to the process, but the dynamics of the fungi are not explicitly considered. Explicit consideration of the dynamics of the fungi (changes in density and composition) allows a third potential way in which these microorganisms can alter plant species coexistence through indirect feedbacks on plant populations (Bever *et al.* 1997). This process builds on the well established observation that plant species differ in their response to individual fungal species. As a result, the composition of the mycorrhizal fungal community can have strong direct effects on the outcome of plant–plant interactions, as is repeatedly demonstrated in manipulative experiments (e.g. van der Heijden *et al.* 1998, Vogelsang *et al.* 2006). A number of studies have shown that plants can benefit other plants indirectly through their support of local symbiotic fungal populations and an established mycorrhizal mycelium in the soil (Kytoviita *et al.* 2003, van der Heijden 2004, Nara & Hogetsu 2004). Studies claiming to show CMNs probably actually reflect altered densities of mycorrhizas (Bever *et al.* 2010).

Whatever the specific mechanism involved, there is unequivocal evidence that some soil fungi can influence, often decisively, plant competitive interactions, at least in some conditions, and that interactions among those fungi can determine the outcome of competition (Hodge & Fitter 2013). Several studies have tested for effects

of soil mycorrhizal fungal diversity on plant community composition (e.g. van der Heijden *et al.* 1998, Wagg *et al.* 2011). Fungal effects have been incorporated into plant community dynamics using ideas of niche modification and plant–soil community feedbacks (Bever *et al.* 1997, 2010, van der Putten *et al.* 2013, Bardgett & van der Putten 2014). These studies show that belowground diversity can influence plant community diversity in multiple ways, which points to the myriad of mechanisms by which complex soil communities impact plant growth, and the potential for differential effects of soil biota to cancel one another out (Wardle *et al.* 2004). Indeed, effects of soil biodiversity on vegetation dynamics operate through a variety of biotic interactions, which influence plant performance and vegetation dynamics directly, through altered herbivory, symbiosis, or pathogenesis, or indirectly through changing soil nutrient availability, predation on the plant-feeding organisms or symbionts, or changing interactions between plants and their aboveground multitrophic communities (Wardle *et al.* 2004, Bezemer & van Dam 2005). In the short term, these biotic interactions can change the capacity of plant species to compete, facilitate, and reproduce, whereas longer-term effects influence fitness and evolutionary adaptation (Bardgett & van der Putten 2014).

An area that is especially rich in new discoveries concerns the role of plant secondary metabolites and defence signals in regulating belowground– aboveground interactions (Biere & Bennett 2013, Soler *et al.* 2012). It was recently discovered that belowground hyphal networks of arbuscular mycorrhizal fungi act as a conduit for defence signals from plants attacked by herbivorous insects (aphids) to adjacent non-attacked plants, thereby acting as an early warning system for herbivore attack (Babikova *et al.* 2013). Also, foliar and shoot herbivory has been shown to exert a unique soil legacy effect which greatly influences the production of defence chemicals in succeeding plants, and that this legacy effect is mediated by alterations in soil fungal community composition (Kostenko *et al.* 2012). These studies illustrate that soil fungi can impact plant growth by modifying biotic interactions between plants and their natural enemies, but the role of soil biodiversity in these processes remains unresolved (Bardgett & van der Putten 2014).

1.2.2 Saprotrophic fungi

Saprotrophy is a nutritional mode which occurs when non-living organic material, other than that killed by the fungus itself, is utilized as a main carbon source.

Plant biomass is the most abundant source of carbon on earth and saprotrophic fungi are considered as the principal agents responsible for its degradation in terrestrial and freshwater habitats. As such they fulfil a vital role in the terrestrial carbon cycle and in humification. In the absence of a decomposition activity of soil organisms, much of the world's land surface would be literally covered

with meters of organic debris (Brussaard *et al.* 2007). Besides carbon cycling, plant litter decomposition represents also the primary route through which mineral nutrients (e.g., N, P, S, Mg...) return to the soil (Berg *et al.* 2001). Therefore, decay of plant organic matter not only controls the balance between soil carbon storage and CO_2 release into the atmosphere, but also the release of essential mineral nutrients, which are again made available for plant growth.

1.2.2.1 Plant litter composition

Plant biomass varies widely in composition depending on plant species, tissue, season and geographical location. It is dominated by polymers, essentially polysaccharides and lignin. Cellulose may constitute 10 - 50% of the litter mass. It is the least complex polymer with a linear structure of β-1,4-linked D-glucose residues. The long glucose chains are tightly bundled together in microfibrils and are non-covalently linked together by hemicelluloses (Kolpak & Blackwell 1976, Carpita & Gibeaut 1993).

Hemicelluloses make up as much as 20-40% of the plant litter (Berg & McClaugherty 2008). They are often heteropolymers derived from several monosaccharides, essentially glucose, xylose, mannose, galactose and arabinose. Hemicelluloses consist of shorter chains (between 70 - 200 sugar units) which are linear or branched.

After cellulose, lignin is the second most abundant biopolymer (Boerjan *et al.* 2003) and comprises 15-40% of litter mass (Berg & McClaugherty 2008). Lignins are heterogeneous and complex polymers with phenolic monomers (p-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol) linked together by several different linkages and most of them are not readily hydrolysable. Lignin is insoluble in water, difficult to hydrolyse and therefore highly resistant to microbial degradation in comparison with polysaccharides and other biopolymers. Lignin is closely associated to cellulose and hemicelluloses in plant cell walls, these compounds together form a complex referred to as lignocellulose (Evans & Hedger 2001).

In addition to cellulose, hemicelluloses and lignin, plant litter also contains in variable proportions other compounds including pectins, tannins, cutin or proteins. Fresh litter also contains low-molecular-weight substances, such as amino acids, simple sugars and short-chain fatty acids.

1.2.2.2 Plant organic matter decomposition by fungi

Biochemical decomposition of plant litter is a sequential process that initially involves the loss of the less recalcitrant components (for example, oligosaccharides, organic acids, hemicellulose and cellulose) followed by the degradation of the remaining highly recalcitrant compounds (for example, lignin or suberin). However,

in the case of substrates like wood, prior degradation of lignin, which physically protects holocellulose, may be required.

About half of terrestrial plant production is holocellulose (cellulose+ hemicellulose), which makes cellulolysis the principle carbon acquisition pathway for decomposer communities. In the early stages of plant litter decomposition, rates of cellulolysis are controlled to a large extent by the availability of nitrogen whereas in the latter stages lignin is rate-limiting. The abundance and composition of lignin is a primary control because it restricts the access of enzymes to cell wall polysaccharides (Talbot *et al.* 2012). In addition, the concentration and composition of lignin and other hydrocarbons (e.g. waxes) in plant litter is thought to influence the quantity of carbon transferred from litter to soil organic matter and ultimately soil carbon sequestration (Theuerl & Buscot 2010).

Saprotrophic fungi have the ability to decompose the recalcitrant lignocellulose fraction of terrestrial organic matter thanks to a combination of morphological characteristics (hyphal growth form) allowing penetration of solid material and long-distance transfer of nutrients through hyphal connections and networks., and physiological characteristics (extracellular enzyme production) (Baldrian & Valásková 2008, Floudas *et al.* 2012). Cellulose and hemicellulose can serve as a sole source of carbon and energy. Lignin is a poor source of energy and polysaccharides serve as a co-substrate for its decomposition (Kirk *et al.* 1976). Decomposition of lignin allows access to cellulose for enzymes.

Lignocellulose degradation by saprotrophic fungi is performed using non-enzymatic and/or enzymatic mechanisms. In the latter and best described case fungi excrete a mixture of oxidative and hydrolytic enzymes whose composition varies between taxa. This enzyme distribution is at the basis for the traditional functional classification of wood-degrading fungi as white rot (predominantly Basidiomycota), brown rot (Ascomycota and Basidiomycota), and soft rot (Ascomycota). These descriptors represent a crude gradient in the relative capacities of the fungi to degrade and mineralize lignin in their quest to obtain carbohydrates for growth (Higuchi 1990, Rabinovich *et al.* 2004; Baldrian 2006, Hoegger *et al.* 2006).

White rot fungi have the enzymatic capacity for complete, rapid breakdown of lignin, although partial or slow break down may be accomplished by other organisms, including certain bacteria (Bugg *et al.* 2011, Brown & Chang 2014). This ability to completely decompose lignin, appears to be mainly restricted to Basidiomycota (Agaricomycotina) (Baldrian 2008, Floudas *et al.* 2012), although lignin breakdown has been reported for the Xylariales, within *the* Ascomycota (Worrall *et al.* 1997, Osono *et al.* 2011 a,b).

White rot fungi use hydrolases that gradually degrade cellulose while lignin is completely mineralized, by contrast brown rot fungi rapidly depolymerise cellulose via oxidative (non-enzymatic) mechanisms, whereas lignin is modified (also *via* non-

enzymatic modification) and remains as a modified polymeric residue without significant mass loss (Blanchette 1995, Yelle *et al.* 2008).

However, a recent comparative genomic analysis (Riley et al. 2014) showed that some fungal species (e.g. the basidiomycetes *Botryobasidium botryosum* and *Jaapia argillacea*) lack ligninolytic enzymes, and thus resemble brown-rot fungi, but possess the cellulose-degrading apparatus typical of white-rot fungi. Moreover, these fungal species appear to degrade lignin, based on decay analyses on wood wafers. These results indicate that the prevailing paradigm of white rot *vs* brown rot does not capture the diversity of fungal wood decay mechanisms, suggesting a continuum rather than a dichotomy between the two modes of wood decay.

Finally, certain ascomycetes fungi can cause soft attack on wood. These fungi are widespread in nature and attack a variety of wood substrates. Soft-rot fungi show preference for cellulose and hemicelluloses; they differ from other decaying fungi by the decay patterns. Some soft-rot fungi produce cavities within the secondary walls of wood cells (Type I attack), while others may erode the secondary wall completely (Type II attack).

1.2.2.3 Enzymes used by fungi for decomposition

Due to the heterogeneity of plant polymers, a large variety of enzymes is needed to degrade them into monomer or short oligomers, which are ultimately used as carbon sources for the fungi (Makela *et al.*, 2014).

These enzymes have been classified in families based on homologies between their amino acid sequences as presented in the Carbohydrates Active enZymes (CAZy) database (www.cazy.org) (Lombard *et al.* 2014), which has become an indispensable resource in the field. CAZy classify proteins as being either glycoside hydrolases (GH), glycosyltransferases (GT), polysaccharide lyases (PL), carbohydrate esterases (CE), auxiliary activities (AA) or carbohydrate binding modules (CBM). Within each of these groups several families can be found, and a subset of these contain enzymes involved in plant biomass degradation. While some families contain only a single (known) enzyme activity, others contain enzymes with different catalytic activities and therefore a family name does not represent an immediate indication of enzyme function. Finally, the number of CAZy entries without functional confirmation by far outnumbers the biochemically characterized entries.

Enzymatic degradation of cellulose can be taken as an example of the complexity of the process and of the enzyme classification system. From an enzymatic point of view, three types of hydrolytic enzymes are usually used by fungi for cellulose degradation : endo-1,4-β-glucanase (EC 3.2.1.4) which cuts within the chains, thus decreasing polymer length and increasing the concentration of polymer extremities, cellobiohydrolase (EC 3.2.1.91) which attack at the end of the chains predominantly producing cellobiose and 1,4-β-glucosidase (EC 3.2.1.21) which

hydrolyse cellobiose to produce two glucose molecules (Baldrian 2008, Baldrian & Valásková 2008). If we take the endo-1,4-β-glucanase activity as an example, enzymes presenting this activity can be found in different CAZy families such as the GH5, GH7 or GH45 ones. While all characterized GH45 have an endo-1,4-β-glucanase activity, members of the GH7 family can either be endo-1,4-β-glucanases or cellobiohydrolases and members of the GH5 can attack various polymers, not only cellulose, but also xylan, mannan or chitosan.

The fungal enzymatic system for lignin degradation in based on oxidative enzymes (AA) with wide substrate specificity. Ligninolytic set of enzymes is composed of oxidases, peroxidases and enzymes producing hydrogen peroxide. Laccases (EC 1.10.3.2) are copper-containing oxidases catalyzing oxidation of phenolic compounds using molecular oxygen as a co-substrate. These enzymes are found in many fungal taxa (Baldrian 2006), including fungi not known to hydrolyse lignin. Class II peroxidases are secreted by several groups of basidiomycetous fungi (Hatakka 1994); based on of conserved catalytic domains and Mn-binding sites they are classified as lignin (EC 1.11.1.14), manganese (EC 1.11.1.13), or versatile peroxidases (EC 1.11.1.16), (Kirk & Farrell 1987, Hammel & Cullen 2008, Martínez 2002).They catalyze oxidation of wide variety of aromatic macromolecules (including lignin and its related compounds). In addition, accessory enzymes such as glyoxalate oxidase (EC 1.2.3.5), glucose-1-oxidase (EC 1.1.3.4) and aryl alcohol oxidase (EC 1.1.3.7) generating hydrogen peroxide required by peroxidases have been found to be involved in lignin degradation (Martínez *et al.* 2005).

Lignocellulose can be also degraded by a non enzymatic system (Fenton based reaction), where hydroxyl radicals are produce and attack lignocellulose molecules, that leads into cellulose and hemicellulose degradation and modification of lignin molecule. The non-enzymatic system is important for lignocellulose degradation by brown rot fungi (Arantes *et al.* 2011).

1.2.2.4 Fungal succession in litter decomposition

It is well known that fungal community composition changes with time during the decomposition of complex organic matter such as litter and wood (Rayner & Boddy 1988, Frankland 1998, Dighton 2007, Osono 2007, Lindahl & Boberg 2008). A major driver of fungal succession in litter is thought to be its chemical composition and in particular the content and chemical structure of lignin (Osono 2007). Initial studies based on the isolation of fungi suggested a succession from endophytes and "primary saprotrophs", mostly ascomycetes, that first used simple sugars and decomposed the easily available cellulose fraction of litter to "secondary decomposers", mostly basidiomycetes that attack lignin (Frankland 1998, Dighton 2007). Similar shifts in fungal community composition have been observed during decay of wood (Rayner & Boddy 1988, Olsson *et al.* 2011). Succession apparently also occurs within the lignin

decomposer community (secondary decomposers), as ascomycetes with limited ligninolytic activity appear to dominate during the early stage of litter decomposition followed by ligninolytic basidiomycetes (Osono 2007). However the general view that lignin is not degraded during the early stage of decomposition (Berg & McClaugherty 2008) has recently been questioned (Koide *et al.* 2005, Osono & Hirose 2009, Klotzbücher *et al.* 2011)

The presence of specific fungi can have a strong impact on the fungal community composition during the succession. Predecessor-successor relationship can be described as "priority effects" as the predecessor creates conditions that have different (positive or negative) effects on the colonization abilities of potential successor species (Fukami *et al.* 2010). Besides the apparent consistent temporal shifts in functional group of fungi during decomposition, a certain degree of specialization of fungal community toward the decomposition of different litter and wood type is also apparent (Osono 2007). The selection of specific decomposer by certain litter types has been proposed to result in the so called "home field-advantage": the presence of best decomposer organisms in soil for a certain litter type as a result of legacies of previously decomposed litter of the same type (Gholz *et al.* 2000).

1.2.2.5 Mycorrhizal fungi and soil organic matter decomposition

Mycorrhizal fungi are direct and indirect actors of soil organic matter decomposition and turn over.

Regarding their direct effects, although mycorrhizal fungi are conventionally regarded as symbionts, their ability to degrade organic matter, and thus act as saprotrophs, has long been discussed (Smith & Read 2008, Baldrian 2009, Cullings & Courty 2009, Talbot *et al.* 2013). Several studies have reported the presence of genes encoding for plant cell wall degrading enzymes in the genomes of ECM fungi (Bödeker *et al.* 2009,2014; Luis *et al.* 2005, Martin *et al.* 2008, Bell *et al.* 2009, Kohler *et al.* 2015) although the global number of such genes is extremely reduced compared to the corresponding numbers in the genomes of true saprotrophs (Kohler *et al.*, 2015). Furthermore, the presence of typical genes encoding hydrolytic enzymes in a genome does necessarily means ability of decomposing the corresponding substrate as in the case of cellulose or xylan for the ECM species *Hebeloma cylindrosporum* (R. Marmeisse, personal communication).

The "mycorrhizal decomposition theory" was hypothesized by Frank already in 1894, in one of the earliest articles on the ectomycorrhizal symbiosis where he stated '...there is no doubt that mycorrhizal fungi account for a major fraction of litter turnover , in order to supply nutrients to the trees'. (*see* Lindahl & Tunlid 2014). Laboratory studies showed that ECM fungi have some ability to decompose various compounds and to express activities of extracellular enzymes thought to be involved

in degradation of plant litter. However ECM fungi have a lower ability to perform decomposition than saprotrophic fungi (Read *et al.* 2005); by contrast ericoid mycorrhizal fungi were proposed to be more efficient decomposers.

Regarding the use by ECM fungi of soil organic matter as a source of nutrients two theories have been proposed. The first one proposed that mycorrhizal fungi benefit from organic matter decomposition as a source of reduced carbon (C) compounds to support metabolism as in the case of true saprotrophs. Such theory is supported by measurements of hydrolytic activities produced by ECM root tips collected in forest soils (Buée *et al.* 2005, Courty *et al.* 2007) and was recently conforted by a report describing strong hydrolytic activities in soil compartments in which only hyphae from ECM fungi had access (Phillips *et al.* 2014). Such "facultative saprotrophism" could enable mycorrhizal fungi to survive in the absence of host plants or when host plants fail to supply enough readily accessible C (Buée *et al.* 2005, Courty *et al.* 2007, Lindahl & Tunlid 2014). Facultative saprotrophism would imply that there is no sharp distintion between the two functional groups, and fungal species would best be described along a biotrophy saprotrophy continuum (Koide *et al.* 2008).

The second theory proposes that ECM fungi "modify" organic matter and more specifically lignin to get access to organic nitrogen sources locked in plant litter. According to this theory, ECM fungi, although interacting with plant cell wall polymers do not however (and cannot) use them as C sources. This theory gained experimental support in a series of experiment using the ECM fungus *Paxillus involutus* (Rineau *et al.* 2012, 2013). Chemical and spectroscopic measurements demonstrated that this fungus was indeed capable of modifying the structure of complex organic molecules without however significantly hydrolysing them (Rineau *et al.* 2012). Furthermore, it was shown that this degradation and access to organic N forms needed glucose (provided in the field by the host plant and therefore could not be equated to saprotrophism (Rineau *et al.* 2013, reviewed in Lindahl & Tunlid 2014)

Regarding the indirect effect of ECM fungi on soil plant organic matter decomposition, it has been shown that ecosystems dominated by ECM plants store more carbon per unit of soil volume than other terrestrial ecosystems (Averill *et al.* 2014). A likely mechanistic explanation is that the plant supply of high-energy monosaccharides to ECM symbionts gives them a strong competitive advantage for the use of organic N forms at the expense of saprotrophic fungal species. As a result, soil saprotrophic fungi are strongly N-limited in ecosystems dominated by ECM plants and therefore do not efficiently degrade dead plant organic matter (Averill *et al.* 2014). An additional indirect support to this theory was provided by Lindahl *et al.* (2010) who showed that severing tree roots (*i.e.* stopping the immediate flux of plant photosynthetates to ECM fungi) resulted in the very short term (5-14 days) in a

significant increase in soil hydrolytic activities (cellulose, laccase) and in an increase in the relative abundance of saprotrophic *versus* ECM taxa.

1.3 Natural diversity as a source for biotechnological application

Soil fungi potentially represent a rich source of novel natural products, including biocatalysts. In addition to their ecological importance, thanks to their capacity to secrete an arsenal of powerful enzymes and compounds, fungi are also major contributors to important both ancient and modern biotechnological processes. Processes and products that utilize fungi include baking, wine making or brewing. Several, easy to grow in pure culture, filamentous fungi are widely used for the production of food processing enzymes and metabolites such as antibiotics and organic acids (Arora *et al.* 2003).

Organic matter decomposition process is also directly related to the process of second generation biofuel production. Plant biomass derived from agriculture, industry and forest (not competing with food production) has indeed been recognized as an alternative and renewable source of energy for the biofuels production. The use of lignocellulolisic materials in energy production can provide environmental (reduction in green house gas emission), economic, and strategic benefits. For this reasons the U.S.A. and the European Union have set ambitious goals with a scenario for supplying 30% and 10% (respectively) of the gasoline demand with biofuels by the year 2030 and 2020 (Viikari *et al.* 2012). Conversion of the cellulosic components into fermentable sugars is, however, still the major technological and economical bottleneck in the production of biofuels or other high-volume commodity products from cellulosic biomass.

Among the aspect that need optimization there are the optimization of the cost of enzymes utilized as well as the optimization their properties (e.g the capacity of binding to the substrates or their thermal stability) and the use of synergic and accessory enzymes (Viikari *et al.* 2012). The discovery of new enzymes or the enzymes modification by genetic- engineering can be powerful tools to overcome those problems. A fundamental contribution can be the one offered by meta-omics approaches, which can give access to the wide biochemical biodiversity of the microorganism performing organic matter decomposition in natural environment (Baldrian & López-Mondéjar 2014).

Soon after the initial proof of concepts (Rondon *et al.* 2000), metagenomics has been proposed as a new way to explore microbial diversity for the discovery of novel biocatalysts and other small molecules of interest (eg Voget *et al.* 2003). To this aim, DNA extracted from environmental samples, which encompass the genomes of the numerous and often non cultivable microorganisms, is cloned in a bacterial expression vector (plasmid, cosmid, phosmid) to constitute environmental metagenomic libraries. These libraries are then transferred in a "domesticated"

bacterial host (*E. coli*, *Pseudomonas sp.*, *Streptomyces sp.*...) and the recombinant colonies screened for specific activities/phenotypes (Simon & Daniel 2011). This has led in the recent years to the description of a wide range of novel enzymes and small molecules of potential interest (e.g. Chistoserdova 2010, Simon & Daniel 2011).

This protocol however excludes eukaryotic genes for several reasons. Eukaryotic DNA is generally less abundant than bacterial DNA in most ecosystems and furthermore eukaryotic genomes are also usually for bigger than bacterial ones. For these reasons selection of a specific eukaryotic gene in a metagenome would require a more intensive effort compared to the selection of a bacterial one. Furthermore, eukaryotic genes are often interrupted by introns and their transcription signals are not recognized by bacterial cells. As an alternative, for eukarya, it has been proposed to use environmental RNA instead of DNA, for the exploration of their functional diversity (Grant *et al.* 2006, Bailly *et al.* 2007). Eukaryotic-specific intron-less polyadenylated mRNA can indeed be selectively isolated from the total environmental RNA, converted into cDNA which can be cloned directionally in a suitable expression vector such as a bacteria/yeast shuttle expression plasmid (Bailly *et al.* 2007). As for metagenomic libraries, metatranscriptomic ones can then be screened by expression in a suitable host. Although this strategy has seldom been followed, its initial implementation has nevertheless demonstrated its potential (Bailly *et al.* 2007, Kellner *et al.* 2011, Todaka *et al.* 2007) and has led to the functional characterization of novel protein families (Damon *et al.* 2012, Lehembre *et al.* 2013).

1.4 Soil fungi and global changes

Due to the pivotal roles fungi play in terrestrial ecosystem functioning, there is a wide interest in addressing the effects of global change on these soil microorganisms.

Research to date suggests that a wide variety of responses may take place in above- or below-ground communities in response to any given change in either community (Wardle *et al.* 2004). Given the multiple interactions of fungi with either living and dead plants, changes in soil fungal community structure may have an impact on aboveground evolutionary processes, including patterns of natural selection on plant traits and plant responses to environmental change (Bardgett & van der Putten 2014). There is a huge body of historic literature reporting how rhizosphere microbes have an impact on plant traits related to nutrient acquisition, drought tolerance, and disease resistance, and ultimately plant fitness, although few studies have been done in non-managed ecosystems (Philippot *et al.* 2013). Recent research also shows that modification of soil microbial communities can impact selection on plant traits with, for example, drought-adapted microbial communities increasing plant fitness under this stress (Lau & Lennon 2012). Similar specificity in selective advantage is exemplified by the finding that litter decomposition can be

more rapid in soil beneath the host plant species, compared to when beneath a different plant species, the so-called home-field advantage (Ayres *et al.* 2009). Home-field advantage effects are not always found and when they are, their strength is highly variable and context dependent. However, recent synthesis suggests that home-field effects are strongest when the quality of 'home' and 'away' litters become more dissimilar, and hence that dissimilarity in plant communities and litter quality between the 'home' and 'away' locations are the most significant drivers of home-field effects (Veen *et al.* 2014). The mechanisms involved in these various community responses still need to be resolved, but it is evident that soil fungal diversity has the potential to impact both evolutionary and ecological processes under global change through direct effects (e.g. of mycorrhizal symbionts), as well as by indirect effects involving decomposer organisms in the soil.

Conversely, effects of global changes on fungi can be direct or indirect (as consequences of effects on organisms they interact with, including plants). Therefore, effects of global change on fungal distribution and activity are hard to predict because they are mediated in many different ways, including: fungal physiology, reproduction and survival, host physiology, spatial and temporal distribution of hosts and resource availability, and outcome of competitive interspecific interactions (Boddy *et al.* 2014). Our understanding of how soil fungi adapt to rapid changes in their environment, whether they can do this fast enough to cope with novel environments, and how this adaptive capacity may relate to the level of soil biodiversity, is limited. A key challenge, therefore, is to determine how soil species respond to rapid environmental change, either through phenotypic plasticity, range shifts or by evolutionary adaptation, how these changes impact aboveground community re-organization and ecosystem functioning, and how the level of soil fungal diversity may influence these processes (Bardgett & van der Putten 2014).

Although scant, evidence is emerging that certain soil fungi have the capacity to respond rapidly to climate change. For instance, an analysis of temporal trends in fungal fruiting patterns in southern England between 1950 and 2005, revealed that climate change has advanced the first and extended the last fruiting date of many fungal species, with probable consequences for decomposition processes in soil (Gange *et al.* 2007). Similarly, an analysis of herbarium records in Norway has revealed that the time of fruiting of mushrooms has changed considerably over recent years, although changes differ across taxa (Kauserud *et al.* 2008). Annual fruiting season of both ECM and saprotrophic fungi was widening during the period 1970-2007 in Austria, Norway, Switzerland and the United Kingdom, the mean annual day of fruiting has become late, but mycorrhizal fungi generally have a more compressed season than saprotrophs, possibly due to the fact that fruiting of mycorrhizal fungi is partly depending on signals from the host plant (Kauserud *et al.* 2012). Changes in fruiting seem to be more similar among species of the same genus

than from species from different genera (Kauserud *et al.* 2012). This extension of the European fungal fruiting season parallels an extended vegetation season in Europe. Whether or not these fungal responses were due to plasticity or evolutionary adaptation has not been established. However, it was recently shown that individual species of decomposer fungi can acclimate to climate change, with warm-acclimated fungi reducing their growth and respiration following warming (Crowther & Bradford 2013).

Global scenarios of biodiversity change in terrestrial biomes for the year 2100 (Sala *et al.* 2000) have identified land use change as the driver that is expected to have the largest global impact on biodiversity, followed by climate change (e. g. elevated temperature and altered precipitation regimes), nitrogen deposition, biotic exchange (the deliberate or accidental introduction of plants and animals to an ecosystem) and changes in atmospheric CO_2 concentration (elevated concentration). In particular, the Mediterranean biome and grassland ecosystems will likely experience the highest proportional change as compared with other biomes because of the considerable influence of all these drivers (Sala *et al.* 2000).

Table 2. Main global change drivers

Global change driver	Scenario/Cause
Land use	Conversion to agriculture
N deposition	Increase over the last 60 years of reactive inorganic nitrogen availability primarily due to increased fertilized usage and fossil fuel consumption (Galloway *et al.* 2003)
Elevated CO2	Atmospheric concentration of CO_2 increased from 1995 to 2005 at a rate of ~2ppm per yr, and is projected to further increase by 40-100% though 2030 (IPCC 2007).
Tropospheric ozone (O_3)	Caused by photochemical reactions involving volatile organic compounds and nitrogen oxides, projections suggest that tropospheric ozone levels will rise as human activity and urbanization increase (Logan, 1985; IPCC, 2007).
Ultraviolet (UV) radiation	Increasing cause by stratospheric ozone depletion
Warming	Global average surface temperature has increased by 0.85°C since 1880, due primarily to an anthropogenic increase in greenhouse gas concentrations, and is predicted to increase by at least 1.5°C by the end of the 21st century (IPCC 2013).
Altered precipitation- drought	Alterations of the quantity, frequency and intensity of precipitation (Kreuzwieser & Gessler 2010; Trenberth 2011). Increasing in episodic rainfall and longer dry periods (IPCC 2007, Kharin *et al.* 2007).

Practices associated with land use change result in unfavourable changes in soil variables (Amundson 2001, Dunjo *et al.* 2003). Therefore, changes in land use and management, such as conversion of temperate grassland into croplands or tropical forests into grasslands, can impact massively on the overall soil community. Mycorrhizal fungal community composition is often altered by land and use transition, in particular increases in land use intensity affect the structure of fungal communities especially in agricultural lands (e.g. Spurgeon *et al.* 2013, Moora *et al.* 2014, Verbruggen *et al.* 2014, Xiang *et al.* 2014).

Mohan *et al.* (2014) recently tried to summarize the responses of mycorrhizal fungi (abundance and activity) to global changes and their influence on ecosystems (productivity, biogeochemical cycling) in a meta-analysis of 120 published papers. The authors emphasize that the effect of global changes were often variable among biomes and fungal species. For instance, mycorrhizal root colonization in response to elevated Nitrogen (N) deposition depends on type of mycorrhizas. ECM and AM are thought to respond differently to increased N availability (Lilleskov 2005). In fact ECM are generally considered to be more sensitive to increase in inorganic N since this group is more "specialized" in plant nitrogen uptake compared to AM fungi (Read 1991), while AM abundance is not consistently affected by increased N availability (Lilleskov 2005). Nitrogen is one of the major factors contributing to the decline in diversity of ECM fungi sporocarps over broad regions of Europe (Lilleskov *et al.* 2001), and experiment and gradient studies suggest that sites with long term N input are also losing diversity belowground (Lilleskov 2002).

Changes in precipitation directly influence soil moisture and fungal communities in soil can directly respond to changes in soil moisture, in fact when water is limiting, constraints on substrate diffusion (Skopp *et al.* 1990) may force hyphal networks to expand. Expansion may also occur for mycorrhizal networks during drought, mycorrhizas are known to improve plant drought tolerance partly through increased rates of water movement from soil into host plants (Augé 2001). However, current evidence for fungal responses to soil moisture is equivocal. For example within individual sites, mycorrhizal hyphae in soils and roots have been shown to increase, decrease, or remain unchanged in response to drought (Miller *et al.* 1995, Lutgen *et al.* 2003; Staddon *et al.* 2003; Clark *et al.* 2009; Querejeta *et al.* 2009).

García-Palacios *et al.* (2014) have identified critical areas for future research related to the effects of global change drivers on soil biota. These areas include soil community composition, altered precipitation, and litter decomposition. The authors underlined indeed that, although the responses of soil biota to global change include abundance, compositional, and physiological shifts (Eisenhauer *et al.* 2012, Wall *et al.* 2012, Frey *et al.* 2013), abundance measurements are most consistently used across taxa and studies (Treseder 2004, 2008, Blankinship *et al.* 2011). Altered

precipitation studies are also underrepresented (<10% of total studies). This constitutes a major gap to understand how soil biota modulates ecosystem responses to global change, because altered precipitation has a larger influence on soil biota abundance across taxa than elevated CO_2 or warming, as found by a recent review on the topic (Blankinship *et al.* 2011). Regarding the ecosystem processes measured, the current underrepresentation of studies assessing litter decomposition complicates the understanding of how soil biota mediates global change effects on nutrient dynamics and C cycling (García-Palacios *et al.* 2014).

1.5 Aim of the thesis

The literature reported in the previous paragraphs illustrates some of the key roles played by soil fungi in essential ecosystems processes, such as soil nutrient mobilization and soil organic matter degradation. It is now clear that soil fungal diversity affects multiple ecosystem processes. However, questions remain over the relative roles of fungal species and functional diversity in driving these processes, and the role of extrinsic factors in modulating fungal diversity and diversity–function relationships (Bardgett & van der Putten 2014).

Research and syntheses over the past 10 years have made it clear that both the identity and the diversity of organisms jointly control the functioning of ecosystems (Cardinale *et al.* 2012). One of the reasons why biodiversity insures ecosystem processes is "because many species provide greater guarantees that some will maintain functioning even if others fail" (Yachi & Loreau 1999). Indeed, it has been proposed that functional redundancy in communities is an insurance for the maintenance of ecosystem processes in case of perturbations causing local species extinctions, which would be compensated by the presence of species that are functionally similar but differ in their responses to changes in environmental factors or disturbances (Walker 1992, Naeem 1998, Yachi & Loreau 1999, Elmqvist *et al.* 2003). Functional redundancy is therefore a critical property for the resilience (i.e. the ability of a community to return its effects on ecosystem processes to a previous state after changing due to a disturbance) of natural communities (Walker 1992, Naeem 1998), which has been demonstrated experimentally (e.g. Joner *et al.* 2011). A recent study testing test causal models linking plant diversity to community stability (Pillar *et al.* 2013) also supports the conclusion that functional redundancy enhances community resilience, therefore corroborating the insurance hypothesis.

The fact that functional redundancy enhances resilience is particularly important for land-use regulation and ecosystem management, given that redundancy tends to decrease with land use intensity (Laliberté *et al.* 2010).

Understanding the factors driving soil fungal diversity is therefore of primary importance to predict the responses of soil fungal communities to global changes and

consequently the impact that these changes may have on ecosystems services. However, hardly anything is known about how soil fungal communities acclimate and adapt to rapid environmental change. Because fungi mainly reside hidden as mycelia belowground or within substrata, it is difficult to monitor changes in their growth and activity within terrestrial ecosystems. Environmental genomics tools, such as meta-barcoding and meta-transcriptomics can help deciphering spatio-temporal changes of soil fungal communities, and therefore such approaches should be implemented to get a better understanding of responses of these organisms to global changes.

In this context, the general aim of the thesis was to assess the impact of two main recognized drivers of global changes on soil fungal diversity, by means of environmental genomics approaches.

I first assessed (*Chapter II*) the impact of changes in land use on one important group of plant symbiotic fungi, the arbuscular mycorrhizal fungi, (AMF). A metabarcoding approach was applied to unravel AMF soil community diversity and responses to different land use practices. This research was conducted in different geographic sites across Europe, as part of the European project "EcoFINDERS: Ecological Function and Biodiversity Indicators in European Soils" (http://www.EcoFINDERS.eu), 2011-2014. The strategic goal of this project was to support European Union soil policy making by providing the necessary tools to design and implement strategies for the sustainable use of soils.

In another chapter (*Chapter III*), I focussed on climate change and more specifically on the alteration in precipitation regime on the diversity of fungi that decompose plant-derived organic matter. The study was conducted in an experimental Mediterranean forest, an ecosystem that is predicted to be strongly impacted by climate changes by the end of the 21st century. To specifically assess changes in the active decomposer communities, an original targeted-metatrascriptomics approach has been implemented. It consisted in the amplification from soil mRNA and systematic sequencing of different fungal genes encoding enzymes active on the main plant cell wall polymers. This approach presents the advantage of providing information about fungi carrying out a specific activity at a specific time. In other words it gave us direct access to the physiologically active saprotrophic fungal community.

Working on fungal genes encoding enzymes involved in plant biomass decomposition in soil brought my interest to the development and application of a new method that could facilitate the recovery of such functional genes from soil samples. *Chapter IV* illustrates how an original modification of the "solution hybrid sequence capture" technique allows efficient recovery of full-length and functional

soil cDNAs encoding enzymes involved in the degradation of plant litter. In addition to its prospective use to study microbial gene diversity in ecosystems, this method could also be used in environmental biotechnology for the bio-prospecting of novel biocatalysts from natural microbial communities.

1.6 References

Amelung, W., Lobe, I., Du Preez, C.C. (2002). Fate of microbial residues in sandy soils of the South African Highveld as influenced by prolonged arable cropping. *Eur. J. Soil Sci.* 53:29–35.

Amundson, R. (2001). The carbon budget in soils. *Annual Review of Earth and Planetary Sciences* 29: 535-562.

Andrews, J.A., Schlesinger, W.H. (2001). Soil CO_2 dynamics, acidification, and chemical weathering in a temperate forest with experimental CO_2 enrichment. *Global Biogeochem. Cycles* 15:149–162.

Arantes, V., Milagres, A.M.F., Filley, T.R., Goodell, B. (2011). Lignocellulosic polysaccharides and lignin degradation by wood decay fungi: The relevance of nonenzymatic Fenton-based reactions. *J. Ind. Microbiol. Biotechnol*. 38:541–555.

Arora, D., Bridge, P., Bhatnagar, D., eds. (2003). Fungal Biotechnology in Agricultural, Food, and Environmental Applications. CRC Press.

Augé, R. (2001). Water relations, drought and vesicular-arbuscular mycorrhizal symbiosis. *Mycorrhiza* 11: 3–42.

Averill, C., Turner, B.L., Finzi, A.C. (2014). Mycorrhiza-mediated competition between plants and decomposers drives soil carbon storage. *Nature* 505:543–5.

Ayres, E. Steltzer, H., Simmons,B.L., Simpson, R.T., Steinweg, J.M, Wallenstein, M.D., Mellor, N. Parton, W.J., Moore,J.C, Wall, D.H. (2009). Home-field advantage accelerates leaf litter decomposition in forests. *Soil Biol. Biochem.* 41: 606–610.

Babikova, Z., Gilbert, L., Bruce, T.J.A., Birkett, M., Caulfield, J.C., Woodcock, C., *et al.* (2013). Underground signals carried through common mycelial networks warn neighbouring plants of aphid attack. *Ecol. Lett.* 16:835–843.

Bailly, J., Fraissinet-Tachet, L., Verner, MC, Debaud, JC, Lemaire, M., Wésolowski-Louvel, M., *et al.* (2007). Soil eukaryotic functional diversity, a metatranscriptomic approach. *ISME J.* 1:632–42.

Baldrian, P. (2006). Fungal laccases-occurrence and properties. *FEMS Microbiol. Rev.* 30:215–242.

Baldrian P. (2008). Wood-inhabiting ligninolytic basidiomycetes in soils: Ecology and constraints for applicability in bioremediation. *Fungal Ecol.* 1:4–12.

Baldrian, P., Valásková, V. (2008). Degradation of cellulose by basidiomycetous fungi. *FEMS Microbiol. Rev*. 32:501–21.

Baldrian, P. (2009). Ectomycorrhizal fungi and their enzymes in soils: Is there enough evidence for their role as facultative soil saprotrophs? *Oecologia* 161:657–660.

Baldrian, P., López-Mondéjar, R. (2014). Microbial genomics, transcriptomics and proteomics: New discoveries in decomposition research using complementary methods. *Appl. Microbiol. Biotechnol.* 98:1531–1537.

Bardgett, R. (2005). The biology of soil. Oxford University press, New York.

Bardgett, R. D., van der Putten, W. H. (2014). Belowground biodiversity and ecosystem functioning. *Nature* 515: 505–511.

Bass, D., Richards, T. (2011). Three reasons to re-evaluate fungal diversity 'on Earth and in the ocean'. *Fungal Biology Reviews* 25: 159-164

Beed, F., Benedetti, A., Cardinali, G., Chakraborty, S., Dubois, T., Garrett, K., *et al.* (2011). Climate change and micro-organism genetic resources for food and agriculture: state of knowledge, risks and opportunities, FAO

Bell, C.W., Acosta-Martinez, V., McIntyre, N.E., Cox, S., Tissue, D.T., Zak, J.C. (2009). Linking microbial community structure and function to seasonal differences in soil moisture and temperature in a Chihuahuan desert grassland. *Microb. Ecol.* 58:827–42.

Berg, B., McClaugherty, C., Santo, AVD, Johnson, D. (2001). Humus buildup in boreal forests: effects of litter fall and its N concentration. *Can J Forest Res* 31: 988–998.

Berg, B., McClaugherty, C. (2008). Plant Litter. Decomposition, humus formation, carbon Sequestration. 2th edition, Springer, Berlin.

Bever, J.D. Westover, K.M., Antonovics, A. (1997). Incorporating the soil community into plant population dynamics: the utility of the feedback approach. *J. Ecol*. 85:561–573.

Bezemer, T. M.,van Dam, N. M. (2005). Linking aboveground and belowground interactions via induced plant defenses. *Trends Ecol. Evol.* 20:617–624.

Biere, A. ,Bennett, A. E. (2013)Three-way interactions between plants, microbes and insects. *Funct. Ecol.* 27: 567–573.

Bever, J.D., Dickie, I.A., Facelli, E., Facelli, J.M., Klironomos, J., Moora, M., *et al.* (2010). Rooting theories of plant community ecology in microbial interactions. *Trends in Ecology & Evolution* 25(8):468–478.

Bidartondo, M. I. (2005). The evolutionary ecology of myco-heterotrophy. *New Phytol.* 2:335–352.

Blackwell, M. (2011). The fungi: 1, 2, 3 ... 5.1 million species? *Am. J. Bot.* 98:426–38.

Blanchette, R.A. (1995). Degradation of the lignocellulose complex in wood. *Can. J. Bot*. 73:999–1010.

Blankinship, J.C., Niklaus, P.A., Hungate, B.A. (2011) A meta-analysis of responses of soil biota to global change. *Oecologa* 165: 553–565.

Boddy, L.,Büntgen, ULF, Egli, S., Gange, A.C., Heegaard, E., Kirk, P.M.,A. Mohammad, Kauserud, H. (2014). Climate variation effects on fungal fruiting. *Fungal Ecology* 10:20-33.

Bödeker, I.T.M., Clemmensen, K.E., de Boer W., Martin F., Olson, Å., Lindahl, B.D. (2014). Ectomycorrhizal Cortinarius species participate in enzymatic oxidation of humus in northern forest ecosystems. *New Phytol.* 203:245–256.

Bödeker, I.T.M., Nygren, C.M.R., Taylor, A.F.S., Olson, A., Lindahl, B.D. (2009). ClassII peroxidase-encoding genes are present in a phylogenetically wide range of ectomycorrhizal fungi. *ISME J.* 3:1387–95.

Boerjan, W., Ralph, J., Baucher, M. (2003). Lignin biosynthesis. *Annu. Rev. Plant Biol.* 54:519–546.

Brovkin, V., van Bodegom, P.M., Kleinen, T., Wirth, C., Cornwell, W.K., Cornelissen, J.H.C., Kattge, J. (2012). Plant-driven variation in decomposition rates improves projections of global litter stock distribution. *Biogeosciences* 9: 565–576.

Brown, M.E., Chang, M.C.Y. (2014). Exploring bacterial lignin degradation. *Curr. Opin. Chem. Biol*. 19:1–7.

Brundrett, M.C. (2002). Coevolution of roots and mycorrhizas of land plants. *New Phytol.* 154:275–304.

Brundrett, M.. (2004). Diversity and classification of mycorrhizal associations. *Biol. Rev. Camb. Philos. Soc.* 79:473–495.

Bruns, T.D., Shefferson, R.P. (2004). Evolutionary studies of ectomycorrhizal fungi: recent advances and future directions. *Can. J. Bot.* 82:1122–1132.

Brussaard, L, de Ruiter, P.C., Brown, G.G. (2007). Soil biodiversity for agricultural sustainability. *Agric. Ecosyst. Environ*. 121:233–244.

Buée, M., Vairelles, D., Garbaye, J. (2005). Year-round monitoring of diversity and potential metabolic activity of the ectomycorrhizal community in a beech (*Fagus silvatica*) forest subjected to two thinning regimes. *Mycorrhiza* 15:235–245.

Buée, M., Reich, M., Murat, C., Morin, E., Nilsson, R.H., Uroz, S., *et al.* (2009). 454 Pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. *New Phytol.* 184:449–56.

Bugg, T.D.H., Ahmad, M., Hardiman, E.M., Singh, R. (2011). The emerging role for bacteria in lignin degradation and bio-product formation. *Curr. Opin. Biotechnol.* 22:394–400.

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P. *et al*. (2012). Biodiversity loss and its impact on humanity. *Nature* 486(7401): 59–67.

Carpita ,N.C., Gibeaut, D.M. (1993). Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J*. 3:1–30.

Chemidlin Prévost-Bouré, N., Dequiedt, S., Thioulouse, J., Lelièvre, M., Saby, N.P.A., *et al.* (2014) Similar processes but different environmental filters for soil bacterial and fungal community composition turnover on a broad spatial scale. *PLoS ONE* 9(11):e111667.

Chistoserdova, L. (2010). Recent progress and new challenges in metagenomics for biotechnology. *Biotechnol. Lett.* 32:1351–1359.

Clark, J.S., Campbell, J.H., Grizzle, H., Acosta-Martìnez, V., Zak ,J.C. (2009). Soil microbial community response to drought and precipitation variability in the Chihuahuan Desert. *Microb. Ecol.* 57:248–60.

Clemmensen, K.E., Bahr, A., Ovaskainen, O., Dahlberg, A., Ekblad, A., Wallander, H., *et al.* (2013). Roots and associated fungi drive long-term carbon sequestration in boreal forest. *Science* 339:1615–8.

Courty, P.E., Bréda, N., Garbaye, J. (2007). Relation between oak tree phenology and the secretion of organic matter degrading enzymes by Lactarius quietus ectomycorrhizas before and during bud break. *Soil Biol. Biochem*. 39:1655–1663.

Crowther, T.W., Bradford, M.A. (2013). Thermal acclimation in widespread heterotrophic soil microbes. *Ecol. Lett*. 16:469–477.

Cullings, K., Courty, P.E. (2009). Saprotrophic capabilities as functional traits to study functional diversity and resilience of ectomycorrhizal community. *Oecologia* 161:661–664.

Daghino, S., Murat, C., Sizzano, E., Girlanda, M., Perotto, S. (2012). Fungal Diversity Is Not Determined by Mineral and Chemical Differences in Serpentine Substrates. *PLoS ONE* 7(9):e44233.

Damon, C., Lehembre, F., Oger-Desfeux, C., Luis, P., Ranger, J., Fraissinet-Tachet L, *et al.* (2012). Metatranscriptomics reveals the diversity of genes expressed by eukaryotes in forest soils. *PLoS One* 7:e28967.

De Vries, F.T., Bloem, J., van Eekeren, N., Brusaard, L., Hoffland, E. (2007). Fungal biomass in pastures increases with age and reduced N input. *Soil Biol. Biochem*. 39:1620–1630.

Dickie, I.A, Xu, B., Koide, R.T. (2002). Vertical niche differentiation of ectomycorrhizal hyphae in soil as shown by T-RFLP analysis. *New Phytol.* 156:527–535.

Dickie, I.A. (2010) Insidious effects of sequencing errors on perceived diversity in molecular surveys. *New Phytol.* 188(4): 916–918.

Dighton,J. (2003). Fungi in Ecosystem Processes., Mycology Series (17), CRC press.

Dighton, J. (2007). Nutrient cycling by saprotrophic fungi in terrestrial habitats. In:*Environmental and microbial relationships*, Vol. 4, Kubicek, C.P., Druzhinina, I.S., eds (ed)., Berlin, Germany, pp. 287–300.

Dominati, E., Patterson, M., Mackay, A. (2010). A framework for classifying and quantifying the natural capital and ecosystem services of soils. *Ecol. Econ*. 69:1858–1868.
Dunjo, G., Pardini, G., Gispert, M. (2003) Land use change effects on abandoned terraced soils in a Mediterranean catchment, NE Spain. *Catena* 52: 23-37.
Egger, K.N., Hibbett, D.S. (2004). The evolutionary implications of exploitation in mycorrhizas. *Can. J. Bot.* 82:1110–1121.
Eisenhauer, N., Cesarz, S., Koller, R., Worm, K., Reich, P.B. (2012) Global change belowground: impacts of elevated CO_2, nitrogen, and summer drought on soil food webs and biodiversity. *Global Change Biology* 18:435–447.
Ekblad A, Näsholm, T. (1996). Determination of chitin in fungi and mycorrhizal roots by an improved HPLC analysis of glucosamine. *Plant Soil* 178:29–35.
Ekblad, A., Wallander, H., Näsholm ,T. (1998). Chitin and ergosterol combined to measure total and living fungal biomass in ectomycorrhizas. *New Phytol.* 138:143–149.
Ekblad, A., Wallander, H., Godbold, D.L., Cruz, C., Johnson, D., Baldrian, P., *et al.* (2013). The production and turnover of extramatrical mycelium of ectomycorrhizal fungi in forest soils: Role in carbon cycling. *Plant Soil* 366:1–27.
Elmqvist, T., Folke, C., Nyström, M., Peterson, G., Bengtsson, J., Walker, B., Norberg, J. (2003). Response diversity, ecosystem change, and resilience. *Frontiers in Ecology and the Environment* 1: 488–494.
Evans, C.S., Hedger, J.N. (2001). Degradation of plant cell wall polymers. In: Fungi in Bioremediation, Gadd, GM, ed (ed)., Cambridge University press: Cambridge, UK, pp. 1–27.
Fitter, A. H., Graves, J. D., Watkins, N. K., Robinson, D., Scrimgeour, C. (1998). Carbon transfer between plants and its control in networks of arbuscular mycorrhizas. *Functional Ecology* 12:406–412.
Floudas, D., Binder, M., Riley, R., Barry, K., Blanchette, RA, Henrissat, B., *et al.* (2012). The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336:1715–9.
Frankland, J.C. (1998). Fungal succession — unravelling the unpredictable. *Mycol. Res.* 102:1–15.
Frey, S.D., Lee, J., Melillo, J.M., Six, J. (2013) The temperature response of soil microbial efficiency and its feedback to climate. *Nature Climate Change* 3:395–398.
Fukami, T., Dickie, I.A., Wilkie P.J., Paulus B.C., Park, D., Roberts, A., *et al.* (2010). Assembly history dictates ecosystem functioning: Evidence from wood decomposer communities. *Ecol. Lett*. 13:675–684.
Galloway, .J.N., Aber, J.D., Erisman, J.W., Seitzinger, S.P., Howarth, R.W., Cowling, E.B., *et al.* (2003). The Nitrogen Cascade. *Bioscience* 53:341.
Gange, A.C., Gange, E.G., Sparks, T.H., Boddy, L. (2007). Rapid and recent changes in fungal fruiting patterns. *Science* 316:71.
García-Palacios, P., Vandegehuchte, M.L., Shaw ,A.E., Dam, M., Post, K.H., Ramirez, K.S., *et al.* (2014). Are there links between responses of soil microbes and ecosystem functioning to elevated CO_2 , N deposition and warming? A global perspective. *Glob. Chang. Biol.* DOI:10.1111/gcb.12788.
Gholz, H.L., Wedin, D.A., Smitherman, S.M., Harmon, M.E., Parton, W.J. (2000). Long-term dynamics of pine and hardwood litter in contrasting environments: Toward a global model of decomposition. *Glob. Chang. Biol.* 6:751–765.
Godbold, D.L., Hoosbeek, M.R, Lukac, M., Cotrufo, M.F., Janssens, I.A., Ceulemans, R., *et al.* (2006). Mycorrhizal hyphal turnover as a dominant process for carbon input into soil organic matter. *Plant Soil* 281:15–24.

Grant, S., Grant ,W.D, Cowan, D.A., Jones, B.E., Ma, Y., Ventosa, A., *et al.* (2006). Identification of eukaryotic open reading frames in metagenomic cDNA libraries made from environmental samples. *Appl. Environ. Microbiol*. 72:135–143.

Hammel, K.E., Cullen D. (2008). Role of fungal peroxidases in biological ligninolysis. *Curr. Opin. Plant Biol.* 11:349–355.

Hatakka, A. (1994). Lignin-modifying enzymes from selected white-rot fungi: production and role from in lignin degradation. *FEMS Microbiol. Rev.* 13:125–135.

Hawksworth, D.L.. (1991). The fungal dimension of biodiversity: magnitude, significance, and conservation. *Mycol. Res.* 95:641–655.

Hawksworth, D.L. (2012). Global species numbers of fungi: Are tropical studies and molecular approaches contributing to a more robust estimate? *Biodivers. Conserv.* 21:2425–2433.

He, X., Critchley, C.,Bledsoe, C. (2003) Nitrogen transfer within and between plantsthrough common mycorrhizal networks (CMNs). *Crit. Rev. Plant Sci.*22:531 –567.

He, X., Critchley, C., Bledsoe, C. (2005). Nodulated N_2- fixing *Casuarina cunninghamiana* is the sink for net N transfer from non-N_2-fixing *Eucalyptus maculata* via an ectomycorrhizal fungus *Pisolithus sp.* using $^{15}NH^{+4}$ or $^{15}NO_3$ supplied as ammonium nitrate. *New Phytol* 167:897–912.

He, X., Bledsoe C., Zasoski, R.J., Southworth, D., Horwath, W.R. (2006). Rapid nitrogen transfer from ectomycorrhizal pines to adjacent ectomycorrhizal and arbuscular mycorrhizal plants in a California oak woodland. *New Phytol* 170:143–151.

van der Heijden, M. G. A., Klironomos. J.N., Ursic, M., Moutoglis, P., Streitwolf-Engel, R., Bolle, T. *et al*. (1998).Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* 396:69–72.

van der Heijden, M.G.A. (2004) Arbuscular mycorrhizal fungi as support systems for seedling establishment in grassland. *Ecol. Lett.* 7: 293 –303.

van der Heijden, M.G.A. (2006) Symbiotic bacteria as a determinant of plant community structure and plant productivity in dune grassland. *FEMS Microbiol. Ecol*. 56:178 –187

van der Heijden, M.G.A., Horton, T.R. (2009) Socialism in soil?The importance of mycorrhizal fungal networks for facilitation in natural ecosystems. *J. Ecol.* 97:1139 –1150

Heinemeyer, A., Wilkinson, M., Vargas, R., Subke, J.A., Casella, E., Morison, J.I.L., *et al.* (2012). Exploring the overflow tap theory: Linking forest soil CO 2 fluxes and individual mycorrhizosphere components to photosynthesis. *Biogeosciences* 9:79–95.

Hibbett, D.S., Taylor, J.W. (2013). Fungal systematics: is a new age of enlightenment at hand? *Nat. Rev. Microbiol*. 11:129–33.

Higuchi, T., (1990). Lignin biochemistry: biosynthesis and biodegradation. *Wood Science and Technology* 24:23-63.

Hobbie, E.A. (2006). Carbon allocation to ectomycorrhizal fungi correlates with belowground allocation in culture studies. *Ecology* 87:563–569.

Hodge, A., Fitter, A. H. (2013). Microbial mediation of plant competition and community structure. *Functional Ecology* 27:865–875.

Hoegger, P.J., Kilaru, S., James, T.Y., Thacker, J.R., Ku, U. (2006) Phylogenetic comparison and classification of laccase and related multicopper oxidase protein sequences. *FEBS J* 273:2308–2326

Högberg, M.N., Briones, M.J.I., Keel, S.G., Metcalfe, D.B., Campbell, C., Midwood, A.J., *et al.* (2010). Quantification of effects of season and nitrogen supply on tree below-ground carbon transfer to ectomycorrhizal fungi and other soil organisms in a boreal pine forest. *New Phytol.* 187:485–493.

Intergovernmental panel on climate change, IPCC. (2007). Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change.

Intergovernmental panel on climate change,IPCC. (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

James, T.Y, Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., Cox, C.J., *et al.* (2006). Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* 443:818–22.

James, T.Y., Berbee, M. L. (2012). No jacket required – new fungal lineage defies dress code: recently described zoosporic fungi lack a cell wall during trophic phase. *Bioessays* 34: 94–102.

James, T.Y., Pelin, A., Bonen, L., Ahrendt, S., Sain, D., Corradi, N., Stajich, J.E. (2013). Shared signatures of parasitism and phylogenomics unite Cryptomycota and Microsporidia. *Curr Biol* 23:1548–1553.

Jeffery, S., Gardi, C., Jones, A., Montanarella, L., Marmo, L., Miko, L., *et al.* (2010). European Atlas of Soil Biodiversity. Publications office of the European Union, Luxembourg.

Johnson, D., Leake, J.R., Ostle, N., Ineson, P., Read, D.J. (2002). In situ $^{13}CO_2$ pulse-labelling of upland grassland demonstrates a rapid pathway of carbon flux from arbuscular mycorrhizal mycelia to the soil. *New Phytol.* 153:327–334.

Johnson, N.C., Graham, J.H., Smith, F.A. (1997). Functioning of mycorrhizal associations along the mutualism-parasitism continuum. *New Phytol*. 135:575–586.

Joner, F., Specht, G., Müller, S.C. Pillar, V.D.(2011). Functional redundancy in a clipping experiment on grassland plant communities. *Oikos* 120: 1420–1426.

Jones, M. D., Forn, I., Gadelha, C., Egan, M.J., Bass, D., Massana, R., Richards, T.A. (2011). Discovery of novel intermediate forms redefines the fungal tree of life. *Nature* 474:200–203.

Karpov, S.A., Mikhailov, K.V., Mirzaeva, G.S., Mirabdullaev, I.M., Mamkaeva, K.A., Titova, N.N., *et al.* (2013). Obligately phagotrophic aphelids turned out to branch with the earliest-diverging fungi. *Protist* 164:195–205.

Karpov, S.A., Mamkaeva, M.A., Benzerara, K., Moreira, D., López-García, P. (2014). Molecular Phylogeny and Ultrastructure of *Aphelidium aff. melosirae* (Aphelida, Opisthosporidia). *Protist* 165:512–526.

Kauserud, H., Stige, L.C., Vik, J.O., Okland, R.H., Høiland, K., Stenseth NC. (2008). Mushroom fruiting and climate change. *Proc. Natl. Acad. Sci. U. S. A*. 105:3811–3814.

Kauserud, H., Heegaard, E., Buntgen, U., Halvorsen, R., Egli, S., Senn-Irlet, B, *et al.* (2012). Warming-induced shift in European mushroom fruiting phenology. *Proc. Natl. Acad. Sci. U. S*. A. 109(36):14488–14493..

Keeling, P.J. (2003). Congruent evidence from alpha -tubulin and beta-tubulin gene phylogenies for a zygomycete origin of microsporidia. *Fungal Genet. Biol.* 38:298–309.

Kellner, H., Luis, P., Portetelle, D., Vandenbol, M. (2011). Screening of a soil metatranscriptomic library by functional complementation of *Saccharomyces cerevisiae* mutants. *Microbiol. Res.* 166:360–8.

Kharin, V.V., Zwiers, F.W., Zhang, X., Hegerl, G.C. (2007). Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *J. Clim.* 20:1419–1444.

Kibblewhite, M.G., Ritz, K., Swift, M.J. (2008). Soil health in agricultural systems. *Philos. Trans. R. Soc. Lond. B. Biol. Sci*. 363:685–701.

Kiers, E. T., Rousseau, R. A., West, S. A., Denison, R. F.(2003). Host sanctions and the legume-rhizobia mutualism. *Nature* 425:78–81.

Kiers, E.T., Denison, R.F. (2008). Sanctions, Cooperation, and the Stability of Plant-Rhizosphere Mutualisms. *Annu. Rev. Ecol. Evol. Syst.* 39:215–236.

Kiers, E.T., Duhamel, M., Beesetty, Y., Mensah, J. A., Franken, O., Verbruggen, E., *et al.* (2011). Reciprocal rewards stabilize cooperation in the mycorrhizal symbiosis. *Science* 333: 880–882.

Kirk, T.K., Connors, W.J., Zeikus, J.G. (1976). Requirement for a growth substrate during lignin decomposition by 2 wood-rotting fungi. *Applied and Environmental Microbiology* 32:192-194.

Kirk ,T.K., Farrell, R.L. (1987). Enzymatic "combustion": the microbial degradation of lignin. *Annu. Rev. Microbiol*. 41:465–505.

Klironomos, J.N., Rillig, C.M, Allen, M.F. (1996). Below-ground microbial and microfaunal responses to artemisia tridentata grown under elevated atmospheric CO_2. *Funct. Ecol*. 10:527–534.

Klotzbücher, T., Kaiser, K., Guggenberger, G., Gatzek, C., Kalbitz, K. (2011). A new conceptual model for the fate of lignin in decomposing plant litter. *Ecology* 92:1052–1062.

Kohler, A., Kuo, A., Nagy, L., Morin, E., Barry, K., Buscot, F.,Canbäck, B. *et al.* (2015).Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nature Genetics*, *in press.*

Koide, K., Osono, T., Takeda, H. (2005). Fungal succession and decomposition of *Camellia japonica* leaf litter. *Ecol. Res.* 20:599–609.

Koide, R.T., Sharda, J.N., Herr, J.R., Malcolm, G.M. (2008). Ectomycorrhizal fungi and the biotrophy-saprotrophy continuum. *New Phytol.* 178:230–233.

Kolpak, F.J, Blackwell, J. (1976). Determination of the structure of cellulose II. *Macromolecules* 9:273–278.

Kostenko, O., van de Voorde, T. F. J., Mulder, P. P. J., van der Putten, W. H., Bezemer, T. M. (2012). Legacy effects of aboveground–belowground interactions. *Ecol. Lett.* 15: 813–821.

Kreuzwieser, J, Gessler, A. (2010). Global climate change and tree nutrition: influence of water availability. *Tree Physiol.* 30:1221–34.

Kytoviita, M.M., Vestberg, M., Tuomi, J. (2003) A test of mutual aid in common mycorrhizal networks: established vegetation negates benefit in seedlings. *Ecology* 84: 898 –906.

Laliberté, E., Wells, J.A., DeClerck, F., Metcalfe, D.J., Catterall, C.P., Queiroz, C., Aubin, I., Bonser, S.P., Ding, Y., Fraterrigo, J.M., McNamara, S., *et al*. (2010). Land-use intensification reduces functional redundancy and response diversity in plant communities. *Ecology Letters* 13: 76–86.

Lara, E., Moreira, D., López-García, P. (2010). The environmental clade LKM11 and *Rozella* form the deepest branching clade of Fungi. *Protist* 161:116–121.

Lau, J.A., Lennon, J.T. (2012)Rapid responses of soil microorganisms improve plant fitness in novel environments. *Proc. Natl Acad. Sci. USA* 109:14058–14062.

Leake, J.R. (1994) The biology of mycoheterotrophic ('saprophytic') plants. *New Phytol* 127:171–216.

Leake, J., Johnson, D., Donnelly, D., Muckle, G., Boddy, L., Read, D. (2004a). Networks of power and influence: the role of mycorrhizal mycelium in controlling plant communities and agroecosystem functioning. *Can. J. Bot*. 82:1016–1045.

Leake, J. R., McKendrick, S. L., Bidartondo, M., Read, D. J. (2004b). Symbiotic germination and development of the myco-heterotroph *Monotropa hypopitys* in nature and its requirement for locally distributed *Tricholoma spp*. *New Phytol* 163:405–423.

Leake, J.R., Ostle, N.J, Rangel-Castro, J.I., Johnson, D. (2006). Carbon fluxes from plants through soil organisms determined by field $^{13}CO_2$ pulse-labelling in an upland grassland. *Appl. Soil Ecol.* 33:152–175.

Lee, S.C., Corradi, N., Byrnes, E.J., Torres-Martinez, S., Dietrich, F.S., Keeling, P.J, *et al.* (2008). Microsporidia evolved from ancestral sexual Fungi. *Curr. Biol.* 18:1675–1679.

Lee, S.C., Corradi, N., Doan, S., Dietrich, F.S., Keeling, P.J., Heitman, J. (2010). Evolution of the sex-related locus and genomic features shared in microsporidia and fungi. *PLoS One* 5(5): e10539.

Lehembre, F., Doillon, D., David, E., Perrotto, S., Baude, J., Foulon, J., *et al.* (2013). Soil metatranscriptomics for mining eukaryotic heavy metal resistance genes. *Environ. Microbiol*. 15:2829–2840.

Lentendu, G., Zinger, L., Manel, S., Coissac, E., Choler, P., Geremia, R. a., Melodelima, C. (2011). Assessment of soil fungal diversity in different alpine tundra habitats by means of pyrosequencing. *Fungal Diversity* 49(1):113–123.

Li, X., Brune, A. (2005a). Digestion of microbial biomass, structural polysaccharides, and protein by the humivorous larva of Pachnoda ephippiata (*Coleoptera: Scarabaeidae*). *Soil Biol. Biochem*. 37:107–116.

Li, X., Brune,A. (2005b). Selective digestion of the peptide and polysaccharide components of synthetic humic acids by the humivorous larva of *Pachnoda ephippiata* (Coleoptera: Scarabaeidae). *Soil Biol. Biochem.* 37:1476–1483.

Lilleskov, E.A., Fahey, T.J., Lovett, G.M. (2001). Ectomycorrhizal fungal aboveground community change over an atmospheric nitrogen deposition gradient. *Ecol. Appl.* 11:397–410.

Lilleskov, E.A, Fahey, T.J., Horton, T.R., Lovett, G.M. (2002). Belowground ectomycorrhizal fungal community change over a nitrogen deposition gradient in alaska. *Ecology* 83:104–115.

Lilleskov, E. (2005). How do composition, structure, and function of mycorrhizal fungal communities respond to nitrogen deposition and ozone exposure? In:*The Fungal community its organization and role in the ecosystem*, Dighton, J., White, J., Oudemans, P., eds (ed). Mycology Vol., CRC Press, pp. 769–801.

Lindahl, B., Boberg, J. (2008). Distribution and function of litter basidiomycetes in coniferous forests. In:*Ecology of Saprotrophic Basidiomycetes*, Boddy, L., Frankland, J.,Van West, eds (ed)., Academic Press: Amsterdam, pp. 183–196.

Lindahl, B.D., de Boer, W., Finlay, R.D. (2010). Disruption of root carbon transport into forest humus stimulates fungal opportunists at the expense of mycorrhizal fungi. *ISME J.* 4:872–881.

Lindahl, B.D., Nilsson, R.H., Tedersoo, L., Abarenkov, K., Carlsen, T., Kjøller, R., *et al*. (2013). Fungal community analysis by high-throughput sequencing of amplified markers - a user's guide. *New Phytol* 199:288–299.

Lindahl, B.D., Tunlid, A. (2014). Tansley insight Ectomycorrhizal fungi – potential organic matter decomposers , yet not saprotrophs. *New Phytol* 205(4):1443-1447.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 42:D490–5.

Luis, P., Kellner, H., Zimdars, B., Langer, U., Martin, F., Buscot, F. (2005). Patchiness and spatial distribution of laccase genes of ectomycorrhizal, saprotrophic, and unknown basidiomycetes in the upper horizons of a mixed forest cambisol. *Microb. Ecol.* 50:570–579.

Lutgen, E.R., Muir-Clairmont, D., Graham, J., Rillig, M.C. (2003). Seasonality of arbuscular mycorrhizal hyphae and glomalin in a western Montana grassland. *Plant Soil* 257:71–83.

Makela, M.R. (2014). Degardation adn modification of plant biomass by fungi. In:*The Mycota V. Fungal Genomics*., Nowrousian, M., ed (ed)., Springer- Verlag: Berlin, pp. 175–208.

Makkonen, M., Matty, P., Berg, I., Handa,T., Hättenschwiler, S.,Ruijven, J., van Bodegom, P.M., Aerts R. (2012). Highly consistent effects of plant litter identity and functional traits on decomposition across a latitudinal gradient. *Ecology Letters* 15(9): 1033–1041.

Martin, J.P., Haider, K. (1979). Biodegradation of ^{14}C-labeled model and cornstalk lignins, phenols, model phenolase humic polymers, and fungal melanins as influenced by a readily available carbon source and soil. *Appl. Environ. Microbiol*. 38:283-289.

Martin, F., Aerts, A., Ahrén, D., Brun, A., Danchin, E.G.J., Duchaussoy, F., *et al.* (2008). The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* 452:88–92.

Martínez, A.T. (2002). Molecular biology and structure-function of lignin-degrading heme peroxidases. *Enzyme Microb. Technol.* 30:425–444.

Martínez, A.T., Speranza, M., Ruiz-Dueñas ,F.J., Ferreira, P., Camarero, S., Guillén, F., *et al.* (2005). Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin.*International Microbiology* 8: 195–204.

Millenium Ecosystem Assessment (MEA). (2003). *Ecosystem and human well-being: a framework for assessment*. Washington, DC: Island Press.

Miller, R.M., Jastrow ,J.D., Reinhardt, D.R. (1995). External hyphal production of vesicular-arbuscular mycorrhizal fungi in pasture and tallgrass prairie communities. *Oecologia* 103:17–23.

Mohan, J.E., Clark J.S., Schlesinger, W.H. (2007). Long-term CO_2 enrichment of a forest ecosystem: implications for forest regeneration and succession. *Ecol. Appl.* 17:1198–212.

Mohan, J.E., Cowden, C.C., Baas, P., Dawadi, A., Frankson, P.T., Helmick, K., *et al.* (2014). Mycorrhizal fungi mediation of terrestrial ecosystem responses to global change: mini-review. *Fungal Ecol.* 10:3–19.

Money, N.P. (2013). Against the naming of fungi. *Fungal Biol.* 117:463–465.

Moora, M., Davison, J., Opik, M., Metsis, M., Saks, U., Jairus, T., *et al.* (2014). Anthropogenic land use shapes the composition and phylogenetic structure of soil arbuscular mycorrhizal fungal communities. *FEMS Microbiol. Ecol.* 90:609–21.

Naeem, S. (1998). Species redundancy and ecosystem reliability. *Conservation Biology* 12: 39–45.

Nakano-Hylander, A. Olsson, P.A. (2007) Carbon allocation in mycelia of arbuscular mycorrhizal fungi during colonisation of plant seedlings. *Soil Biol. Biochem.* 39: 1450 – 1458.

Nara, K., Hogetsu, T. (2004) Ectomycorrhizal fungi on established shrubs facilitate subsequent seedling establishment of successional plant species. *Ecology* 85:1700 –1707.

Nylund, J., Wallander, H. (1992). Ergosterol analysis as a means of quantifying mycorrhizal biomass. In:*Methods in microbiology*, Norris, J., Read, D., Varma, A., eds (ed)., Academic Press: London, pp. 77–88.

O'Brien, B.L., Parrent, J.L, Jackson, J.A., Moncalvo, J.M, Vilgalys, R. (2005). Fungal community analysis by large-scale sequencing of enviromental samples. *Appl. Environ. Microbiol* 71: 5544–5550.

Olsson, J., Jonsson, B.G., Hjältén, J., Ericson, L. (2011). Addition of coarse woody debris - The early fungal succession on *Picea abies* logs in managed forests and reserves. *Biol. Conserv*. 144:1100–1110.

Oono, R., Denison, R. F., Kiers, E. T. (2009). Tansley review: controlling the reproductive fate of rhizobia: how universal are legume sanctions? *New Phytol.* 183:967–979.

Oono, R., Anderson, C. G.,. Denison, R. F. (2011). Failure to fix nitrogen (N_2) by nonreproductive symbiotic rhizobia triggers sanctions that reduce fitness of their reproductive clonemates. *Proc. R. Soc. Lond.* 278:2698–2703.

Orgiazzi, A., Lumini, E., Nilsson, R.H., Girlanda, M., Vizzini, A., Bonfante, P., Bianciotto, V. (2012). Unravelling soil fungal communities from different Mediterranean land-use backgrounds. *PLoS One* 7:e34847.

Orgiazzi, A., Bianciotto, V., Bonfante, P., Daghino, S., Ghignone, S., Lazzari, A, Lumini, E., Mello, A., Napoli, C., Perotto, S., Vizzini, A., Bagella, S., Murat, C., Girlanda, M. (2013). 454 pyrosequencing analysis of fungal assemblages from geographically distant, disparate soils reveals spatial patterning and a core mycobiome. *Diversity* 5:73–98.

Orwin, K.H., Kirschbaum, M.U.F., St John, M.G., Dickie, I.A.. (2011). Organic nutrient uptake by mycorrhizal fungi enhances ecosystem carbon storage: A model-based assessment. *Ecol. Lett*. 14:493–502.

Osono, T. (2007). Ecology of ligninolytic fungi associated with leaf litter decomposition. *Ecol. Res*. 22:955–974.

Osono, T., Hirose, D. (2009). Effects of prior decomposition of *Camellia japonica* leaf litter by an endophytic fungus on the subsequent decomposition by fungal colonizers. *Mycoscience* 50:52–55.

Osono, T., Hagiwara, Y., Masuya, H. (2011a). Effects of temperature and litter type on fungal growth and decomposition of leaf litter. *Mycoscience* 52:327–332.

Osono, T., To-Anun, C., Hagiwara, Y., Hirose, D. (2011b). Decomposition of wood, petiole and leaf litter by *Xylaria* species from northern Thailand. *Fungal Ecol*. 4:210–218.

Parrent. J.L., Vilgalys, R. (2007). Biomass and compositional responses of ectomycorrhizal fungal hyphae to elevated CO_2 and nitrogen fertilization. *New Phytol.* 176:164–174.

Pellissier, L., Niculita-Hirzel, H., Dubuis, A., Pagni, M., Guex, N., Ndiribe, C., *et al.*(2014). Soil fungal communities of grasslands are environmentally structured at a regional scale in the Alps. *Molecular Ecology*. 23: 4274–4290.

Pfeffer, P.E., Douds, D.D., B¨ucking, H., Schwartz, D.P., Shachar-Hill, Y. (2004).The fungus does not transfer carbon to or between roots in an arbuscular mycorrhizal symbiosis. *New Phytol* 163: 617-627.

Phillips, L.A., Ward, V., Jones, M.D. (2014). Ectomycorrhizal fungi contribute to soil organic matter cycling in sub-boreal forests. *ISME J.* 8:699–713.

Philippot, L., Raaijmakers, J.M., Lemanceau, P., van der Putten, W.H. (2013)Going back to the roots: the microbial ecology of the rhizosphere. *Nature Rev. Microbiol.* 11: 789–799.

Pietikainen, A. and Kytoviita, M.M. (2007) Defoliation changes mycorrhizal benefit and competitive interactions between seedlings and adult plants. *J. Ecol.* 95:639 –647.

Pillar, V.D., Blanco, C.C., Müller, S.C., Sosinski, E.E., Joner, F., & Duarte, L. D. S. (2013). Functional redundancy and stability in plant communities. *Journal of Vegetation Science*, 24:963–974.

van der Putten, W. H., Bardgett, R. D., Bever, J. D., Bezemer, T. M., Casper, B. B., Fukami, T., *et al*. (2013). Plant-soil feedbacks: The past, the present and future challenges. *Journal of Ecology* 101: 265–276.

Querejeta, J., Egerton-Warburton, L.M., Allen, M.F. (2009). Topographic position modulates the mycorrhizal response of oak trees to interannual rainfall variability. *Ecology* 90:649–662.

Quince, C., Lanzen, A., Curtis, T.P., Davenport, R.J., Hall, N., *et al.* (2009). Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat Methods* 6(9): 639–641.

Rabinovich, M.L., Bolobova, A.V., Vasilchenko, L.G., 2004. Fungal decomposition of natural aromatic structures and xenobiotics: a review. *Applied Biochemistry and Microbiology* 40, 1-17.

Rayner, A.D., Boddy, L. (1988). Fungal decomposition of wood; its biology and ecology. John Wiley Press, Chichester, New York.

Read, D.J. (1991). Mycorrhizal in ecosystems. *Cell. Mol. life Sci.* 47:376–391.

Read, D.J., Leake, J.R., Perez-Moreno J. (2005). Mycorrhizal fungi as drivers of ecosystem processes in heathland and boreal forest biomes. *Can. J. Bot.* 83:1073.

Redecker, D. (2000). Glomalean Fungi from the Ordovician. *Science* 289:1920–1921.
Reeder, J., Knight, R. (2009) The 'rare biosphere': a reality check. *Nat Methods* 6(9): 636–637.
Regus, J. U., Gano, K. A.,. Hollowell, A. C., Sachs, J. L. (2014). Efficiency of partner choice and sanctions in Lotus is not altered by nitrogen fertilization. *Proc. R. Soc. Lond. B* DOI:10.1098/rspb.2013.2587.
Riley, R., Salamov, A.A., Brown,D.W., Nagy, L.G., Floudas, D., Held, B.W, Levasseur, A., *et al.* (2014). Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. *Proc. Natl. Acad. Sci. U. S*. A 111: 9923–9928.
Rineau, F., Roth, D., Shah, F., Smits, M., Johansson, T., Canbäck,B., *et al.* (2012). The ectomycorrhizal fungus *Paxillus involutus* converts organic matter in plant litter using a trimmed brown-rot mechanism involving Fenton chemistry. *Environ. Microbiol.* 14:1477–1487.
Rineau, F., Shah, F., Smits, M.M., Persson, P., Johansson, T., Carleer, R., *et al.* (2013). Carbon availability triggers the decomposition of plant litter and assimilation of nitrogen by an ectomycorrhizal fungus. *ISME J.* 7:2010–22.
Robinson, D., Fitter, A.(1999) The magnitude and control of carbon transfer between plants linked by a common mycorrhizal network. *J Exp Bot* 50:9–13.
Roesch, L.F.W., Fulthorpe, R.R., Riva,A., Casella, G., Hadwin, A.K.M., Kent, A.D., *et al.* (2007). Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J.* 1:283–90.
Rondon, M.R., August, P.R., Bettermann, A.D., Brady, S.F., Grossman, T.H., Liles, M.R., *et al.* (2000). Cloning the soil metagenome: A strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol*. 66:2541–2547.
Rosling, A., Cox, F., Cruz-Martinez, K., Ihrmark, K., Grelet, G.A., Lindahl, B.D. *et al.* (2011). Archaeorhizomycetes: unearthing an ancient class of ubiquitous soil fungi. *Science* 333:876–879.
Sala, O.E., Chapin III, F.S., Armesto, J., Berlow, E., Bloomfield, J., Dirzo, R., *et al.* (2000). Global Biodiversity Scenarios for the Year 2100. *Science* 287:1770–1774.
Schmidt, P.A., Bálint,.M., Greshake, B., Bandow, C., Römbke, J., Schmitt, I. (2013). Illumina metabarcoding of a soil fungal community. *Soil Biol. Biochem*. 65:128–132.
Schüβler, A., Schwarzott, D., Walker, C. (2001). A new fungal phylum, the Glomeromycota: phylogeny and evolution. *Mycol. Res.* 105:1413–1421.
Selosse, M.A., Richard, F., He, X., Simard, S.W. (2006). Mycorrhizal networks: des liaisons dangereuses? *Trends Ecol. Evol.* 21:621–628.
Selosse, M.A., Rousset, F. (2011). The Plant-Fungal Marketplace. *Science* 333:828-829.
Simard, S.W., Perry, D.A., Jones, M.D., Myrold, D.D., Durall, D.M., Molina, R. (1997) Net transfer of carbon between ectomycorrhizal tree species in the field. *Nature* 388:579 –582.
Simard, S.W., Durall, D., Jones, M. (2002) Carbon and nutrient fluxes within and between mycorrhizal plants. In: *Mycorrhizal ecology.* van der Heijden, M.G.A., Sanders, I.R. (eds) Springer, Berlin Heidelberg New York, pp 33–74.
Simard, S.W., Durall, D.M. (2004) Mycorrhizal networks: a review of their extent, function, and importance. *Can J Bot* 82(8):1140–1165.
Simms, E. L., D. L. Taylor, J. Povich, R. P. Shefferson, M. Urbina, Y. Tauszick. (2006). An empirical test of partner choice mechanisms in a legume- rhizobium interaction. *Proc. R. Soc. Lond. B* 273:77–81.
Simon, C., Daniel, R. (2011). Metagenomic analyses: Past and future trends. *Appl. Environ. Microbiol.* 77:1153–1161.
Skopp, J., Jawson, M.D, Doran, J.W. (1990). Steady-state aerobic microbial activity as a function of soil water content. *Soil Sci. Soc. Am. J*. 54:1619.
Smith, S., Read, D. (2008). Mycorrhizal symbiosis. Academic Press.

Soler, R.v der Putten ,W.H., Harvey, J.A., Vet, L.E., Dicke, M., Bezemer, T.M. (2012).Root herbivore effects on aboveground multitrophic interactions: patterns, processes and mechanisms. *J. Chem. Ecol*. 38: 755–767.

Solomon, D., Lehmann, J., Zech, W. (2001). Land use effects on amino sugar signature of chromic Luvisol in the semi-arid part of northern Tanzania. *Biol. Fert. Soils* 33:33-40.

Spurgeon, D.J., Keith, A.M., Schmidt, O., Lammertsma, D.R., Faber, J.H. (2013). Land-use and land-management change: relationships with earthworm and fungi communities and soil structural properties. *BMC Ecol.* 13:46.

Staddon, P.L, Fitter, A.H. (1998). Does elevated atmospheric carbon dioxide affect arbuscular mycorrhizas? *Trends Ecol. Evol.* 13:455–458.

Staddon, P.L., Thompson, K., Jakobsen, I., Grime, J.P., Askew, A.P, Fitter, A.H. (2003). Mycorrhizal fungal abundance is affected by long-term climatic manipulations in the field. *Glob. Chang. Biol*. 9:186–194.

Strickland, M.S., Rousk, J. (2010). Considering fungal:bacterial dominance in soils - Methods, controls, and ecosystem implications. *Soil. Biol. Biochem*. 42:1385-139.

Talbot, J.M., Yelle, D.J., Nowick, J., Treseder, K.K. (2012). Litter decay rates are determined by lignin chemistry. *Biogeochemistry* 108:279–295.

Talbot, J.M., Bruns, T.D., Smith, D.P., Branco,S,. Glassman, S.I. Erlandson, S., *et al.* (2013). Independent roles of ectomycorrhizal and saprotrophic communities in soil organic matter decomposition. *Soil Biol. Biochem*. 57:282–291.

Taylor, D. L., Bruns, T. D., Leake, J. R., Read, D. J. (2002). Mycorrhizal specificity and function in myco-heterotrophic plants. In *Mycorrhizal Ecology,* van der Heijden M. , Sanders, I. (Eds.), pp. 375–413. Springer,Berlin, Heidelberg, NewYork.

Taylor, A. F. S. (2006), Common mycelial networks: life-lines and radical addictions*. New Phytol* 169: 6–8.

Taylor, D.L., Herriott, I. C., Stone, K.E., McFarland, J., Booth, W.M. *et al.* (2010). Structure and resilience of fungal communities in Alaskan boreal forest soils. *Canadian Journal of Forest Research* 40:1288 – 1301.

Taylor, D. L., Hollingsworth, T. N., McFarland, J. W., Lennon, N. J., Nusbaum, C., Ruess, R. W. (2014). A first comprehensive census of fungi in soil reveals both hyperdiversity and fine-scale niche partitioning. *Ecological Monographs* 8:3–20.

Tedersoo, L., Jairus, T., Horton, B.M., Abarenkov, K., Suvi, T., Saar, I., Kõljalg, U. (2008) Strong host preference of ectomycorrhizal fungi in a Tasmanian wet sclerophyll forest as revealed by DNA barcoding and taxon-specific primers. *New Phytol* 180: 479 –490.

Tedersoo, L., Nilsson, R.H., Abarenkov, K., Jairus, T., Sadam, A. *et al.* (2010). 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New Phytol* 188: 291–301.

Teste, F.P.,Simard, S.W., Durall, D.M., Guy, R.D., Berch, S.M. (2010) Net carbon transfer between Pseudotsuga menziesii var. glauca seedlings in the field is influenced by soil disturbance. *J. Ecol.* 98:429 –439.

Theuerl, S., Buscot ,F. (2010). Laccases: Toward disentangling their diversity and functions in relation to soil organic matter cycling. *Biol. Fertil. Soils* 46:215–225.

Tibbett, M. Sanders, F.E. (2002) Ectomycorrhizal symbiosis can enhance plant nutrition through improved access to discrete organic nutrient patches of high resource quality. *Ann. Bot.* 89: 783 –789.

Timling, I., Walker, D. A., Nusbaum, C., Lennon, N. J., Taylor, D. L. (2014). Rich and cold: Diversity, distribution and drivers of fungal communities in patterned-ground ecosystems of the North American Arctic. *Molecular Ecology* 23:3258–3272.

Todaka,N., Moriya, S., Saita, K., Hondo, T., Kiuchi, I., Takasu, H., *et al.* (2007). Environmental cDNA analysis of the genes involved in lignocellulose digestion in the symbiotic protist community of Reticulitermes speratus. *FEMS Microbiol. Ecol.* 59:592–599.

Trenberth, K. (2011). Changes in precipitation with climate change. *Clim. Res*. 47:123–138.

Treseder, K.K. (2004). A meta-analysis of mycorrhizal responses to nitrogen, phosphorus, and atmospheric CO_2 in field studies. *New Phytol*. 164:347–355.

Tuffen, F., Eason, W. R., Scullion, J. (2002). The effect of earthworms and arbuscular mycorrhizal fungi on growth of and ^{32}P transfer between *Allium porrum* plants. *Soil Biology and Biochemistry* 34 (7):1027-1036.

Turbe´, A., De Toni, A., Benito, P., Lavelle, P., Ruiz, N., van der Putten, W., Labouze, E.S.M. (2010). Soil biodiversity: functions, threaths and tools for policy makers. In *Bio Intelligence Service*, IRD, and NIOO, Report for European Commission (DG Environment).

Veen, G.F., Freschet, G.T., Ordonez, A., Wardle, D. A. Litter quality and environmental controls of home-field advantage effects on litter decomposition. *Oikos* doi.org/10.1111/oik.01374

Verbruggen, E., Rillig, M.C., Wehner, J., Hegglin, D., Wittwer, R., van der Heijden, M.G.A. (2014). Sebacinales, but not total root associated fungal communities, are affected by land-use intensity. *New Phytologist*. doi:10.1111/nph.12884

Viikari, L., Vehmaanperä, J., Koivula, A. (2012). Lignocellulosic ethanol: from science to industry. *Biomass and Bioenergy* 46:13–24.

Vogelsang, K. M., Reynolds, H. L., Bever, J. D. (2006). Mycorrhizal fungal identity and richness determine the diversity and productivity of a tallgrass prairie system. *New Phytologis*, 172:554–562.

Voget, S., Leggewie, C., Uesbeck, A., Raasch, C., Jaeger, K.E., Streit, W.R. (2003). Prospecting for novel biocatalysts in a soil metagenome. *Appl. Environ. Microbiol.* 69:6235–6242.

Wagg, C., Jansa, J., Stadler, M., Schmid, B., van der Heijden, M. G. A. (2011). Mycorrhizal fungal identity and diversity relaxes plant–plant competition. *Ecology* 92:1303–1313.

Walker, B.H. (1992). Biodiversity and ecological redundancy. *Conservation Biology* 6: 18–23.

Wall, D.H., Bardgett, R.D., Behan-Pelletier, V. (2012). Soil Ecology and Ecosystem Services. Oxford University Press, Oxford.

Wallander, H., Nilsson, L.O., Hagerberg, D., Baath, E. (2001). Estimation of the biomass and seasonal growth of external mycelium of ectomycorrhizal fungi in the field. *New Phytol.* 151:753–760.

Wallander, H., Ekblad, A., Godbold, D.L., Johnson, D., Bahr, A., Baldrian,P., Björk, R.G. *et al.* (2013). Evaluation of methods to estimate production, biomass and turnover of ectomycorrhizal mycelium in forests soils – A review. . *Soil Biol. Biochem* 57:1034-1047.

Wardle, D.A., Hörnberg, G., Zackrisson, O., Kalela-Brundin, M., Coomes, D.A. (2003). Long term effects of wildfire on ecosystem properties across an island area gradient. *Science* 300: 972-975.

Wardle, D. A., Bardgett, R.D., Klironomos, J.N., Setälä, H., van der Putten, W.H., Wall, D.H. (2004) Ecological linkages between aboveground and belowground biota. *Science* 304:1629–1633.

Wilson, G.W.T., Hartnett, D.C., Rice, C.W. (2006). Mycorrhizal-mediated phosphorus transfer between tallgrass prairie plants Sorghastrum nutans and *Artemisia ludoviciana. Funct. Ecol.* 20:427 –435.

Worrall, J.J., Anagnost, S.E., Zabel, R.A. (1997). Comparison of wood decay among diverse lignicolous Fungi. *Mycologia* 89:199.

Wu, T., Ayres, E., Li, G., Bardgett, R.D., Wall, D.H., Garey, J.R. (2009). Molecular profiling of soil animal diversity in natural ecosystems. *Soil Biol. Biochem*. 41:849-857.

Wyatt, G.A.K., Kiers, E.T., Gardner, A., West, S.A. (2014). A biological market analysis of the plant-mycorrhizal symbiosis. *Evolution* 68(9):2603–2618.

Xiang, D., Verbruggen, E., Hu, Y., Veresoglou, S.D., Rillig, M.C., Zhou, W., *et al.* (2014). Land use influences arbuscular mycorrhizal fungal communities in the farming-pastoral ecotone of northern China. *New Phytol.* 204(4):968-78.

Yelle, D.J., Ralph, J., Lu, F., Hammel, K.E. (2008). Evidence for cleavage of lignin by a brown rot basidiomycete. *Environ. Microbiol*. 10:1844–1849.

Yachi, S., Loreau, M. (1999). Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proc. Natl. Acad. Sci. U. S*. A. 96:1463–1468.

Chapter II

Effect of different land uses on arbuscular mycorrhizal fungal communities and taxa in soil

2.1 Foreword

Land use change is the most important driver of changes in biodiversity (Sala 1995). It has effects on plants species and associated animals and also (most severely) on below-ground organisms (Anderson 1995). Intensive land use necessary for providing food and materials is known to have damaging effect on soil. However, the effects on soil organisms are less well understood. One of the aims of the European project EcoFINDERS (Ecological Function and Biodiversity Indicators in European Soils, http://ecofinders.eu/), started in 2011, was to describe soil biodiversity under different land management regimes in order to link biodiversity to ecosystem functioning and services.

In this context, the biodiversity workpackage (WP1) aimed at analysing the diversity of all biota (microbes, fauna and plants) of representative European soils. The University of Torino research group was involved with the INRA-Dijon research group in the assessment of the diversity of Arbuscular Mycorrhizal Fungi (AMF). I had the opportunity to participate to this workpackage and to carry out the work described in this chapter.

The name of this group of fungi is derived from the *arbuscle*, the typical tree-like structure they form inside the plant root cells, used for nutritional exchanges between the plant and the fungus. AMF in soil can be found as spores and as hyphae forming the extra-radical mycelium compartment, which is connected to the intra-radical mycelium and is fundamental to explore the soil environment and to uptake mineral nutrients.

With the advent of high-throughput sequencing technologies (e.g. 454 pyrosequencing) in 2009, AMF soil assemblages started to be investigated in large-scale studies by means of DNA-metabarcoding strategies. Studies carried out with these strategies highlighted an unexpected diversity of AMF in roots and soils (e.g. Lumini *et al.* 2009, Öpik *et al.* 2009).

Considering the power of DNA-metabarcoding coupled with high- throughput sequencing technologies, the EcoFINDERS project aimed at describing AMF soil communities by means of this approach.

The aim of this work was to investigate the influence of different land use practices on the taxonomic diversity of AMF soil communities as well as on individual AMF taxa.

References:

Anderson, J.M. (1995). Section 6. In:Global Biodiversity Assessment, Mooney, HA, Lubchenco, J, Dirzo, R, & Sala, O.E,, eds (ed)., Cambridge University press: Cambridge, UK, pp. 406–412.

Lumini, E., Orgiazzi ,A., Borriello, R., Bonfante, P., Bianciotto, V.. (2009). Disclosing arbuscular mycorrhizal fungal biodiversity in soil through a land-use gradient using a pyrosequencing approach. *Environ. Microbiol*. 12:2165–2179.

Öpik, M., Metsis, M., Daniell, T.J., Zobel, M., Moora, M. (2009). Large-scale parallel 454 sequencing reveals host ecological group specificity of arbuscular mycorrhizal fungi in a boreonemoral forest. *New Phytol.* 184:424–37.
Sala, O.E. (1995). Section 5. In:Global Biodiversity Assessment, Mooney, HA, Lubchenco, J, Dirzo, R, & Sala, OE (eds), Cambridge University press, Cambridge, UK.

Title: Effect of different land uses on arbuscular mycorrhizal fungal communities and taxa in soil

C. Bragalini[1,2], M.-L. Bouffaud[4,5,6], M. Peyret-Guzzon[4], H. Stockinger[4], D. Van Tuinen[4], S. Voyron[1], E. Lumini[3], D. Wipf[1], D. Redecker[4], V. Bianciotto[3], M. Girlanda[1,3]

Affiliations:
[1]*University of Torino, Italy;*
[2]*Université Claude Bernard Lyon1, France;*
[3]*CNR-Istituto per la Protezione Sostenibile delle Piante, UOS Turin (CNR-IPSP), Italy;*
[4]*INRA/ Université de Bourgogne, UMR 1347 Agroecology, Dijon, France;* [5]*Helmholtz Center for Environmental Research – UFZ, Halle, Germany;*
[6]*German Centre for Integrative Biodiversity Research (iDiv) Halle – Jena – Leipzig, Germany.*

Manuscript to be submitted

2.2 Introduction

The arbuscular mycorrhiza is a widespread mutualistic symbiosis between the majority of land plants (80% of all plant species; Smith & Read 2008) and fungi from the phylum Glomeromycota, originated 450 million years ago (Redecker 2000). Only 250 species of Glomeromycota have been described based on spore morphology, which is a relatively small number considering their global geographic distribution and numerous potential plant hosts.

Arbuscular mycorrhizal fungi (AMF) transfer phosphorus and other mineral nutrients from soil to the plant, thereby improving plant growth, in exchange of host's photosynthates (up to 30%; Drigo *et al.* 2010). They can confer plant pathogen protection (Azcón-Aguilar & Barea 1997, Zhang *et al.* 2009), as well as improve plant tolerance to heavy metals contaminants (Hildebrandt *et al.* 2007) and drought (Augé 2001, Li e*t al.* 2013). Various ecosystems services are linked to the numerous functions provided by this group of fungi, such as increased plant productivity (van der Heijden & Klironomos 1998, Lekberg & Koide 2005), influences on the cycles of carbon, phosphorus and nitrogen (Fitter *et al.* 2011), and maintenance of soil structure and stability (Mummey & Rillig 2006).

Reflecting the important role played by local AMF communities in determining plant growth, and because of the current environmental threats to AMF diversity (Turrini & Giovannetti 2012), there is increasing interest in describing and explaining the distribution of AMF diversity in human-dominated landscapes (Moora *et al.* 2014).

AMF communities are sensitive to a variety of biotic and abiotic factors, including various aspects of land management which modify the physical and chemical characteristics of the soil and the plant cover/composition. A number of studies addressing AMF community structure have indeed shown the influence of the local environment on AMF. Potential determinants of AMF diversity and community composition are the identity of the host plants (e.g. Bainard *et al.* 2011, Liu *et al.* 2012), soil type (Verbruggen *et al.* 2012, Hazard *et al.* 2013), soil pH (An *et al.* 2008, Dumbrell *et al.* 2010a, Hazard *et al.* 2013, Bainard *et al.* 2014, Xiang *et al.* 2014), soil texture and nutrient concentrations (Lekberg *et al.* 2007, Fitzsimons *et al.* 2008, Bai *et al.* 2009, Verbruggen *et al.* 2012, Moebius-Clune *et al.* 2013, Bainard *et al.* 2014, Xiang *et al.* 2014), climate (Dumbrell *et al.* 2011, Hazard *et al.* 2013). Apart from these specific environmental factors, direct land use-related circumstances, such as tillage (which increases soil disturbance) or grazing (which removes aboveground plant biomass), have also been suggested as drivers of AMF community composition (e.g. Helgason *et al.* 1998, Oehl *et al.* 2010, Schnoor *et al.* 2011). As previously underlined (Jansa *et al.* 2014), soil quality, health and management history are extremely important factors for understanding and supporting the sustainable use of soils, but can be difficult to measure directly, and information on these critical aspects in most cases depend on information transmitted by landowners or farmers, with limited independent verification options (to cross-check the validity and/or precision of this information). Due to the relatively low diversity of AMF taxa, the different preferences of the individual taxa for environmental properties, and their global distribution, AMF are promising candidates as bioindicators of land management legacies and soil quality degradation (Jansa *et al.* 2014).

A number of works have therefore addressed the impact of agricultural practices and changes in land use on either the composition of AMF communities (e.g. van der Gast *et al.* 2011, Lin *et al.* 2012, Orgiazzi *et al.* 2012, Verbruggen *et al.* 2012, Dai *et al.* 2013, Hazard *et al.* 2013, Morris *et al.* 2013, Moora *et al.* 2014, Xiang *et al.* 2014) or individual taxa (Bainard *et al.* 2014, Jansa *et al.* 2014). Experimental studies, however, are usually confined to one or a few sites, and may therefore have facilitated the identification of effects of specific agricultural practices at the expense of visibility of other (independent) effects such as soil type and geography (Jansa *et al.* 2014). By contrast, with a couple of recent exceptions (Öpik *et al.* 2013, Tedersoo *et al.* 2014), global factors governing the structuring of AMF communities have been inferred from meta-analyses (e.g. Öpik *et al.* 2006, 2010, Kivlin *et al.* 2011). The latter approach, however, suffers from the limited comparability of individual studies due to the major confounding effects of different sample preparation methods and different experimental methodologies.

Furthermore, to fully appreciate the extent of anthropogenic influence on AMF soil communities, the impact of management practices should be weighed

against the background of the normal (e.g. temporal) fluctuations under field conditions (Pereira e Silva *et al.* 2013).

In this study, soil AMF assemblages were described in different Long-Term Observatories (mainly grasslands) in different European climatic and geological zones, in spring and autumn 2011, by means of 454 pyrosequencing of the internal transcribed spacer 2 (ITS2) rDNA region. The primary goal of this study was therefore to determine (i) what were the main factors that shape AMF community structures and the abundances of individual AMF taxa in different sites in Europe, managed differentially, and (ii) to identify individual AMF taxa or combinations of taxa suitable for use as biomarkers of land use intensification. Since the relative role of the various drivers of AMF community composition may change at different spatial scales (Horn *et al.* 2014), comparative analyses of AMF communities and taxa distributions were performed at both the continental (among-LTO) and local (intra-LTO) scales.

We specifically asked: (1) Does land use intensification has a larger impact compared to other environmental filters in structuring AMF communities at individual field sites and across Europe? (2) Is the impact of land use intensification on AMF communities mediated by changes in soil physico-chemical features? (3) Are there "universal" AMF indicators of land use changes across different geographic and climatic regions?

2.3 Materials and Methods

2.3.1 *Sites and soil sampling*

In this study, four Long-Term Observatories (LTOs) distributed over western Europe, encompassing a range of climatic zones, soil types and land uses, were compared. At each site two levels of land use intensification were considered (low- and high-intensity). The first LTO is situated in Berchidda, Sardinia, Italy (40°49'N, 9°17''E), where two different management types of Mediterranean agricultural lands were analyzed: (i) an Intensive Grassland (IG) which is pastured and mown every 1-5 years for fodder production, and (ii) a Wooded Pasture (WP) dominate by grass species with a low cork-oak tree density. The second LTO is located in Lancaster, United-Kingdom (54°18'N, 2°10'W) and two levels of fertilization on grassland ecosystem were compared: (i) an Improved grassland (I) which is fertilized and has a low floristic diversity, and (ii) an Unimproved grassland (U) with a high floristic diversity, managed traditionally. The third LTO is situated in Lusignan, France (46°24'N, 0°7'E) and includes two cultural practices: (i) a Permanent Culture (PC) of maize and (ii) a Permanent Grassland (PG) with nitrogen amendment. The fourth LTO is located in the National Park of Veluwe, Netherlands (52°03'N, 5°45'E) with two levels of abandoned agricultural lands: (i) a long term abandoned grassland with low human activities (Low), and (ii) a short-term abandoned grassland with high

human activities (High). A minimum of three spatially independent soil samples were sampled from all treatments in spring and autumn 2011. Geographic distances between LTO (inter-LTO distances) were comprised between 597 and 1715 km, and within LTO (between plots with different land uses) were between 0.0664 and 0.400 km. Soil samples of a specific land use were collected at 100 m of distance from each other.

Detailed information on sampling procedure and soil processing are available at https://www.youtube.com/watch?v=_k7BEInBXEc. Soil was sieved at 2mm mesh and stored at -40°C at the GenoSol platform (http://www2.dijon.inra.fr/plateforme_genosol/en) until DNA extraction.

2.3.2 *Soil physicochemical analyses*

Several soil properties were measured by the Laboratoire d'analyse des sols d'Arras of INRA (http://www.lille.inra.fr/las). Total nitrogen (N), total carbon (C) and organic matter contents were measured after combustion at 1000°C. Phosphorus (P) content was determined by $NaHCO_3$ (0.5 M, pH 8.5) extraction (Olsen 1954). The cation exchange capacity (CEC) was determined by extraction with $Co(NH_3)_6Cl_3$ (Ciesielsky & Sterckeman 1997). Exchangeable cations (Ca, Mg, Na, K, Fe, Mn and Al) were extracted using cobaltihexamine and determined by inductively coupled plasma spectrometry-atomic emission spectrometry (ICP- AES). Soil pH was measured on soil slurry (1:5 deionised water:soil).

2.3.3 *DNA extraction and purification*

Total genomic DNA was extracted from 1g of each sample using the ISOm protocol, described in Plassart *et al.* (2012). DNA extracts were purified in two steps. First, DNA was loaded onto polyvinylpolypyrrolidone (PVPP) minicolumns (BIORAD, Marne-la-Coquette, France) and centrifugated at 1,000g 2 min at 10°C. Then, the eluate was purified using the Geneclean turbo kit (Q-Biogene, Illkirch, France). Purified DNA was quantified using the Picogreen kit (Invitrogen, Saint Aubin, France) according to the manufacturer's instructions.

2.3.4 *PCR amplification and pyrosequencing*

Nested PCRs were performed on all samples and each DNA extract was amplified in three replicates. The first PCR was performed using 0.4U of Phusion High Fidelity DNA polymerase (Thermo Fisher Scientific, Courtaboeuf, France), 1x Phusion HF buffer, 0.5µM of the primers SSUmCf and LSUmBr (Krüger et al. 2009), 0.2mM of each dNTPs and 1µl of genomic DNA, in a final volume of 20µl. The PCR conditions used were 5 min at 99°C, 35 cycles of 10 s at 99°C, 30 s at 63°C and 1 min at 72°C, followed by 10 min at 72°C, using an Eppendorf Mastercycler epgradient S (Vaudaux-Eppendorf, Schönenbuch, Switzerland). Each PCR product was checked on agarose

gel, and diluted at 1/50 to use as template in the nested PCR. The nested PCR was done using 1U of Phusion High Fidelity polymerase, 1x HF buffer, 0.5µM of the primers ITS3m (White *et al.* 1990, modified unpublished) and ITS4 (White *et al.* 1990) with barcodes, 0.2µM of each dNTPs and 2µl of diluted PCR product, in a total volume of 50µl. PCR conditions were 30 s at 98°C, 30 cycles of 10 s at 98°C, 30 s at 64°C and 20 s at 72°C, followed by 10 min at 72°C, in an Eppendorf Mastercycler epgradient S. PCR products were checked on agarose gel, the three replicates of each sample were pooled and purified using the High Pure PCR Product Purification Kit (Roche Applied Science, Meylan, France) following the manufacturer's instructions. After quantification with Picogreen, the purified PCR products were mixed equimolarly to prepare sequencing libraries. The libraries were sent to Beckman Coulter Genomics (Grenoble, France) for sequencing using 454 GS FLX technology.

2.3.5 *Sequence and data analysis*

The raw data were split in different fasta files after removing primer and barcode sequences. The filtered sequences were trimmed and denoised using Mothur v.1.30.2 (Schloss *et al.* 2009). The resulting sequences were clustered using Uclust (Edgar 2010) at 97% identity threshold to create Operational Taxonomic Units (OTUs), and singletons were excluded from further analysis. OTUs were used to perform rarefaction analysis with EstimateS software v.9.0.0 (Colwell R. K. 2013, http://viceroy.eeb.uconn.edu/EstimateS/).

For taxonomic assignment, firstly a Blast search against UNITE database was performed in order to eliminate non-Glomeromycota sequences and sequences for which the best Blast hit had a e-value> 1.10^{-5}. Secondly, the EPA algorithm of RAxML (Berger *et al.* 2011) was used to correct and improve the taxonomic assignment of the OTUs. The different datasets obtained were rarefied to the same sequencing depth before performing further analysis.

Jaccard distances between AMF communities (based on presence/absence data matrix) were calculated and visualized in non-metric multidimensional scaling (NMDS) plots. The effects of LTO, season, land use intensity and land use type factors on AMF community composition were evaluated using permutational multivariate analysis of variance (PERMANOVA, 1000 permutations).

The indicators species analysis (Dufrene & Legendre 1997) was carried out, to assess if and which species/combinations of two species were associated with a particular land use type.

Data obtained from basic soil characteristics (e.g. pH, soil granulometry, organic matter content, macronutrient content; see Table 4) measurements were submitted to a principal component analysis (PCA). In addition, soil chemical properties were included in a second matrix in NMDS analysis to identify which

environmental variables were significantly linked to ordination of AMF community (using the *envfit* function of the R vegan package, Oksanen *et al.* 2013).

Correlations between matrices of community distances (Jaccard distances) and matrices of soil properties distances (Euclidean distances) and of geographic distances were analysed by means of Mantel tests based on 999 permutations.

To quantify the relative contribution of soil parameters, land use intensity and season upon soil AMF communities, variance partitioning was performed. Soil variables which had a significant effect on microbial community structure were selected using the *ordistep* function (R vegan package, Oksanen *et al.* 2013). Subsequently, the *varpart* (R vegan package, Oksanen *et al.* 2013) function was used to determine the amount of variance in AMF communities that could be explained by forward selected soil conditions, land use intensity and season.

All the analysis were carried out in R environment with packages *vegan* (Oksanen *et al.* 2013) and *indicspecies* (De Cáceres & Legendre 2009). Plots were drawn using the *ggplot2* package (Wickham 2009).

2.4 Results

This study analyses arbuscular mycorrhizal fungal (AMF) assemblages in the soils of four Long Term Observatories (LTOs) across Europe by means of 454-pyrosequencing of ITS2 amplicons.

Each LTOs was characterized by two types of land use (featuring contrasting levels of intensification -low *vs* high intensity-; Table 1), and soil samples were collected in two seasons (spring and autumn 2011).

2.4.1 *Soil properties*

The analysed LTOs represent a range of soils with different physico-chemical properties (Table 4).

Soil properties did not change significantly between sampling seasons (PERMANOVA; Table 2) and therefore, subsequent analyses were run on the combined (spring + autumn) datasets. Soil properties varied significantly between LTOs and were affected by land use (Tables 2 and 3). For example, soil from the Lusignan LTO was characterized by a higher pH compared to the other LTOs (Figure 1: PCA biplot; Tables 4-5) while the Veluwe one had a higher P content and C/N ratio. Berchidda was the LTO in which the two land uses differed most. Indeed, 12 out of 18 soil parameters differed significantly between the intensive grassland (IG) and the wooded pasture land use (WP, Figure 1: PCA biplot; Table 6). By contrast, in the case of the other LTOs, only two (fine loam and fine sand content for Lusignan) and four (clay, fine sand, P and Fe content for Lancaster and clay, P, Fe and Al content for Veluwe) parameters differed significantly between the two land use types (Table 6).

The intensification levels (low *vs* high) did not have a significant effect on soil properties when the four LTOs were analysed together (Table 2). However, the interaction between the LTO and intensity factors indicated some LTO-specific effect. Indeed, the interaction between the LTO and intensity factors was significant in all pairwise comparisons involving Berchidda (Table 3).

Long Term Observatories (LTOs)	*Lat.*	*Long.*	*Description*	*High intensity land use*	*Low intensity land use*
Lusignan (France)	46° 24' N	0°7' E	2 cultural practices	Permanent culture of maize (PC)	Permanent grassland with nitrogen amended (PG)
Berchidda (Sardinia, Italy)	40°49' N	9°17'' E	2 Mediterranean management types	Intensive Grassland (IG)	Wooded Pasture (WP)
Lancaster (UK)	54°18' N	2°10' W	2 levels of grassland fertilization	Improved grassland (I)	Unimproved grassland (U)
Veluwe (Netherlands)	52°03' N	5° 45'E	2 levels of abandoned agricultural lands	Short term abandoned grassland with high human activities (High)	Long Term abandoned grassland, with human activities (Low)

Table 1 Description of the four Long Term Observatories

Factor	r^2	*p-value*
LTO (n=4)	0.672	**0.000999**
Season (n=2)	0.004	0.987
Intensity (n=2)	0.135	0.6154
Land use (n=8)	0.745	**0.000999**
LTO*season	0.017	0.623377
LTO*intensity	0.060	**0.001998**
season*intensity	0.002	0.9980
season*land use	0.021	0.993007

Table 2. Results of PERMANOVA (1000 permutations) based on the Euclidean distance matrix of the scaled soil parameters. Significant results are in bold type.

	Lusignan vs Berchidda		Lusignan vs Veluwe		Lusignan vs Lancaster		Berchidda vs Veluwe		Berchidda vs Lancaster		Veluwe vs Lancaster	
Factor	r^2	*p-value*	r^2	*p-value*	r^2	*p-value*	r^2	*p-value*	r^2	*p-value*	r^2	*p-value*
LTO	0.571	**0.000999**	0.571	**0.000999**	0.483	**0.000999**	0.487	**0.000999**	0.425	**0.000999**	0.583	**0.0009999**
season	0.024	0.7762	0.031	0.6394	0.012	0.95	0.012	0.9331	0.005	0.993	0.004	0.986
intensity	0.155	**0.03996**	0.059	0.2787	0.031	0.42396	0.063	0.179	0.060	0.1099	0.034	0.3067
land use	0.820	**0.000999**	0.6627	**0.000999**	0.534	**0.000999**	0.686	**0.000999**	0.562	**0.000999**	0.641	**0.000999**
LTO*season	0.021	0.406593	0.034	0.167832	0.023	0.303696	0.009	0.86131	0.017	0.485515	0.002	0.971029
LTO*intensity	0.098	**0.000999**	0.029	0.192807	0.020	0.358649	0.145	**0.000999**	0.0797	**0.004995**	0.023	0.120879
season*intensity	0.008	0.98202	0.00472	0.9980	0.00319	1	0.003	0.9980	0.008	0.9590	0.003	0.998
season*land use	0.030	0.526474	0.04122	0.730270	0.029	0.947053	0.007	0.999001	0.025	0.934066	0.007	0.999

Table 3. Results of PERMANOVA analyses (1000 permutations) on all pairwise comparisons of the four LTOs (Euclidean distance matrices of the scaled soil parameters). Significant results are in bold type.

LTO	*Clay* (< 2 µm) g/Kg	*Fine Loam* (2/20µm) g/Kg	*Coarse Loam* (20/50µm) g/Kg	*Fine Sand* (50/200µm) g/Kg	*Coarse Sand* (200/2000µm) g/Kg	*Organic C* g/Kg	*Total N* g/Kg	*C/N*	*pH*	*P* (P2O5 Olsen)
Berchidda	146.21 *(16.31)*	120.21 *(11.40)*	76.70 *(6.97)*	143.12 *(10.12)*	513.76 *(40.87)*	21.99 *(2.40)*	1.61 *(0.16)*	13.67 *(0.62)*	5.65 *(0.36)*	0.02 *(0.01)*
Lusignan	174.75 *(24.75)*	374.63 *(27.50)*	280.63 *(49.00)*	85.50 *(27.66)*	83.25 *(19.63)*	11.43 *(1.67)*	1.13 *(0.13)*	10.11 *(0.97)*	6.52 *(0.16)*	0.06 *(0.01)*
Veluwe	61.62 *(20.06)*	27.87 *(11.34)*	33.03 *(17.33)*	213.92 *(99.15)*	663.56 *(99.96)*	29.48 *(5.18)*	1.42 *(0.23)*	20.91 *(2.79)*	5.30 *(0.50)*	0.20 *(0.07)*
Lancaster	284.37 *(47.23)*	183.77 *(30.00)*	127.50 *(25.51)*	270.07 *(50.21)*	134.30 *(67.65)*	57.05 *(12.86)*	5.17 *(1.18)*	11.19 *(1.58)*	5.42 *(0.35)*	0.04 *(0.03)*

LTO	*CEC* Cobaltihexamine cmol+/kg	*Ca* cmol+/kg	*Mg* cmol+/kg	*Na* cmol+/kg	*K* cmol+/kg	*Fe* cmol+/kg	*Mn* cmol+/kg	*Al* cmol+/kg
Berchidda	6.92 *(2.96)*	4.89 *(2.49)*	1.28 *(0.72)*	0.14 *(0.04)*	0.21 *(0.07)*	0.01 *(0.01)*	0.07 *(0.02)*	0.54 *(0.41)*
Lusignan	6.61 *(0.69)*	6.02 *(0.44)*	0.22 *(0.19)*	0.03 *(0.01)*	0.17 *(0.09)*	0.01 *(0.00)*	0.25 *(0.18)*	0.05 *(0.01)*
Veluwe	2.98 *(1.12)*	2.94 *(1.27)*	0.29 *(0.09)*	0.03 *(0.01)*	0.10 *(0.05)*	0.01 *(0.01)*	0.02 *(0.01)*	0.50 *(0.33)*
Lancaster	16.13 *(3.96)*	13.87 *(4.6)*	1.97 *(0.66)*	0.15 *(0.05)*	0.27 *(0.17)*	0.01 *(0.00)*	0.18 *(0.11)*	0.45 *(0.38)*

Table 4. Mean values and standard deviations (within brackets) of the soil variables measured at each LTO (combined dataset: spring+autumn).

	all LTOs	Lusignan *vs* Berchidda	Lusignan *vs* Veluwe	Lusignan *vs* Lancaster	Berchidda *vs* Veluwe	Berchidda *vs* Lancaster	Veluwe *vs* Lancaster
Clay	**1,29E-09**	0,3084	**0,001054**	**0,001616**	**0,0002171**	**3,58E-05**	**1,05E-05**
Fine loam	**5,93E-10**	**0,001876**	**0,001101**	**0,0003071**	**0,000224**	**0,0001123**	**1,06E-05**
Coarse loam	**4,00E-10**	**0,001876**	**0,001101**	**0,0003013**	**0,000224**	**3,52E-05**	**1,05E-05**
Fine sand	**6,51E-07**	**0,001876**	**0,009481**	**0,0003071**	**7,01E-01**	**3,58E-05**	**4,07E-01**
Coarse sand	**6,49E-09**	**0,001876**	**0,001101**	0,7972	**0,008414**	**3,58E-05**	**1,06E-05**
Organic carbon	**8,53E-10**	**0,00197**	**0,00115**	**0,0003167**	**0,00864**	**3,70E-05**	**1,10E-05**
Total Nitrogen	**4,45E-09**	**0,00197**	0,1009	**0,0003167**	0,4069	**3,70E-05**	**1,10E-05**
C/N	**2,46E-08**	**0,00197**	**0,00115**	0,84	**0,0002338**	**0,00615**	**1,09E-05**
pH	**3,97E-05**	**0,00197**	**0,00115**	**0,0003159**	0,4932	0,3617	1
Phosphorus	**3,90E-08**	**0,003585**	**0,001145**	**0,01471**	**0,0002311**	0,2317	**1,56E-05**
CEC_cobaltiheamine	**3,20E-09**	1	**0,00115**	**0,0003167**	**0,005752**	**6,59E-05**	**1,10E-05**
Magnesium (Mg)	**2,65E-08**	**0,003665**	**0,0003167**	1	0,2459	**0,0002338**	**1,10E-05**
Sodium (Na)	**4,75E-08**	**0,001946**	1	**0,0003137**	**0,000232**	1	**1,09E-05**
Potassium (K)	**2,44E-05**	1	0,1226	0,1514	**0,003052**	0,7294	**0,0001357**
Iron (Fe)	**0,01767**	0,104	1	1	0,9048	**0,0258**	0,8379
Manganese (Mn)	**4,63E-08**	**0,01541**	**0,001139**	1	**0,000261**	**0,003656**	**1,09E-05**
Alluminium (Al)	**9,94E-05**	**0,00197**	**0,001145**	**0,0003167**	1	1	1
Calcium (Ca)	**1,21E-08**	1	**0,00115**	**0,0003167**	0,1924	**0,000167**	**1,10E-05**

Table 5. Results of non-parametric Kruskal-Wallis test (all LTOs) and Mann-Whitney pairwise comparisons (Bonferroni-corrected p-values) on the 18 soil physico-chemical variables measured at each LTO (combined dataset: spring+autumn). Significant p-values (<0.05) are in bold type.

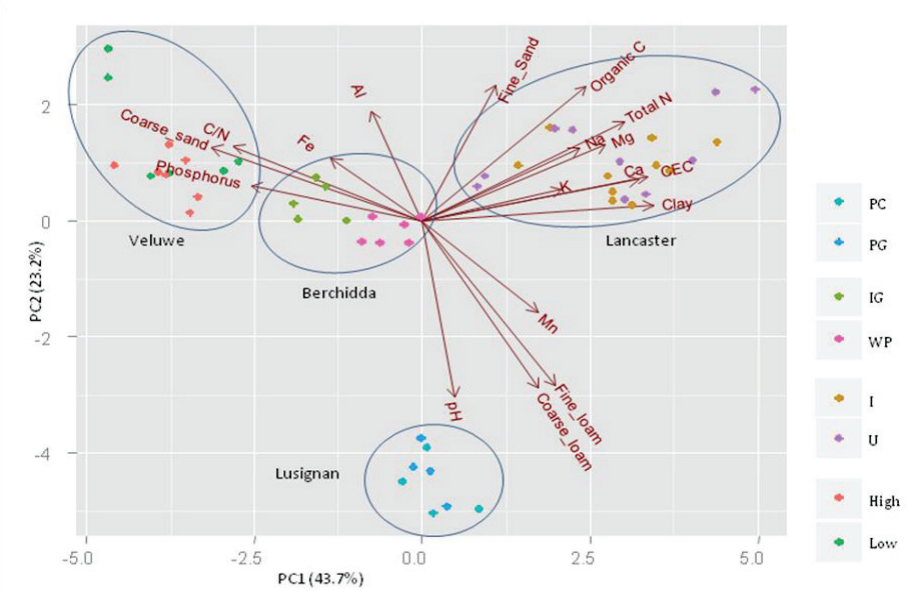

Figure 1. PCA biplot of the soil properties data (combined dataset: spring+autumn). Arrows represent directions and effect of each particular soil variable in discriminating between soils from different LTOs.

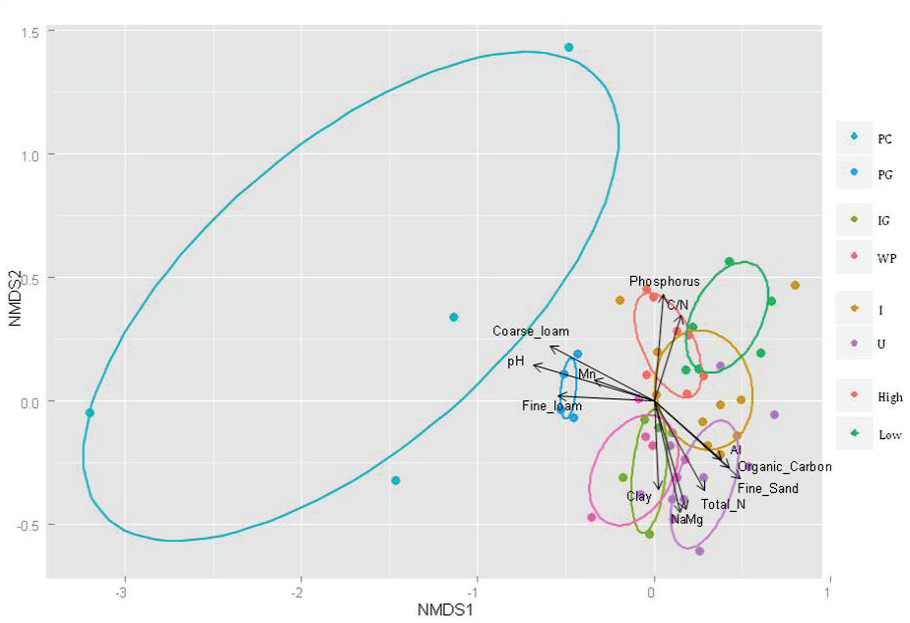

Figure 2. Nonmetric multidimensional scaling (NMDS) ordination of AMF soil communities based on the Jaccard presence-absence distance matrix. Stress of the final NMDS solution was 0.1839577. Arrows correspond to significantly fitted environmental parameters in the ordination ($p<0.05$, Table 10).

	Lusignan		*Berchidda*		*Lancaster*		*Veluwe*	
	PC	*PG*	*IG*	*WP*	*I*	*U*	*High*	*Low*
Clay	178.00 (33.49)	171.50 (16.74)	**159.73** (8.82)	**134.94** (11.59)	**274.20** (30.96)	**294.53** (59.34)	**78.71** (3.46)	**41.67** (8.02)
Fine loam	**353.00** (16.17)	**396.25** (16.01)	**131.33** (5.61)	**110.94** (2.79)	201.13 (25.12)	166.40 (24.47)	29.76 (12.61)	25.67 (10.35)
Coarse loam	305.00 (26.56)	256.25 (57.56)	80.20 (5.89)	73.78 (6.84)	139.40 (28.86)	115.60 (15.04)	35.43 (18.06)	30.22 (17.66)
Fine Sand	**65.50** (6.35)	**105.50** (26.03)	**151.00** (6.09)	**136.56** (7.84)	**239.40** (26.88)	**300.73** (50.10)	174.81 (97.45)	259.56 (86.93)
Coarse sand	97.50 (15.59)	69.00 (10.71)	**477.73** (25.80)	**543.78** (20.72)	145.87 (81.17)	122.73 (52.69)	681.29 (82.45)	642.89 (121.94)
Organic C	10.33 (1.37)	12.52 (1.21)	21.31 (0.69)	22.57 (3.20)	58.03 (9.09)	56.07 (16.26)	31.44 (4.80)	27.20 (5.02)
Total N	1.06 (0.16)	1.20 (0.04)	1.63 (0.07)	1.59 (0.21)	5.19 (0.45)	5.14 (1.65)	1.40 (0.24)	1.45 (0.24)
C/N	9.82 (0.86)	10.41 (1.11)	**13.05** (0.25)	**14.18** (0.18)	11.18 (1.40)	11.21 (1.81)	22.69 (2.16)	18.83 (1.87)
pH	6.62 (0.08)	6.42 (0.16)	**5.32** (0.27)	**5.92** (0.11)	5.35 (0.35)	5.49 (0.36)	5.52 (0.28)	5.06 (0.60)
P	0.07 (0.01)	0.05 (0.01)	**0.03** (0.01)	**0.01** (0.00)	**0.05** (0.03)	**0.02** (0.01)	**0.25** (0.06)	**0.15** (0.05)
CEC	6.73 (0.97)	6.50 (0.36)	**3.98** (0.43)	**9.37** (1.23)	16.01 (3.68)	16.26 (4.41)	3.37 (1.20)	2.53 (0.92)
Ca	6.11 (0.64)	5.93 (0.18)	**2.44** (0.33)	**6.93** (1.16)	13.36 (3.67)	14.39 (5.53)	3.44 (1.20)	2.36 (1.19)
Mg	0.21 (0.22)	0.23 (0.17)	**0.53** (0.08)	**1.89** (0.15)	2.20 (0.64)	1.74 (0.64)	0.31 (0.09)	0.26 (0.10)
Na	0.03 (0.00)	0.03 (0.01)	0.12 (0.03)	0.17 (0.03)	0.16 (0.05)	0.13 (0.03)	0.03 (0.01)	0.02 (0.00)
K	0.24 (0.07)	0.10 (0.01)	0.25 (0.10)	0.18 (0.02)	0.26 (0.06)	0.29 (0.24)	0.12 (0.06)	0.08 (0.04)
Fe	0.01 (0.00)	0.01 (0.00)	**0.02** (0.01)	**0.01** (0.00)	**0.01** **(0.00)**	**0.01** **(0.00)**	**0.01** (0.00)	**0.02** (0.02)
Mn	0.26 (0.23)	0.25 (0.15)	0.07 (0.01)	0.07 (0.02)	0.21 (0.11)	0.15 (0.10)	0.02 (0.01)	0.03 (0.01)
Al	0.05 (0.01)	0.04 (0.01)	**0.96** (0.17)	**0.20** (0.05)	0.45 (0.34)	0.45 (0.44)	**0.35** (0.14)	**0.67** (0.42)

Table 6 Mean values and standard deviations (within brackets) of the soil physico-chemical parameters in high- and low-intensity land uses at the four LTOs (combined dataset: spring+autumn). Bold values indicate significantly different values of a given soil variable between land use treatments at each LTO (Mann-Whitney-U tests, p-value<0.05)

2.4.2 *AMF community structure analyses*

2.4.2.1 *Global comparisons (continental scale)*

The fungal sequences obtained from the spring and autumn samplings were analysed together, yielding a total of 5920 OTUs, defined at 97% identity threshold, including 1023 singletons. Rarefaction analysis indicated saturation of species richness (for each land use, data not show). These OTUs were blasted against the UNITE database, and the taxonomic assignment was improved using the EPA-RAxML algorithm, using known AMF species. A total 324,734 sequences were assigned to 57 Glomeromycota taxa (Table S1), belonging to all Glomeromycota orders (Glomerales, Archaesporales, Diversisporales and Paraglomerales).

To analyse the whole dataset (4 LTOs x 2 intensification levels x 2 seasons), a subsampling of 800 reads per sample was performed before comparisons, and samples with less than 800 sequences were excluded from the analyses. In order to avoid PCR primers bias, we only considered species occurrences (presence/absence data) for β-diversity assessment (Jaccard distances). As for the soil properties, no effect of the season on the AMF community structure was detected (PERMANOVA, Table 7), and therefore subsequent analyses were run exclusively on the combined datasets (spring+autumn).

These analyses highlighted significant effects of both the LTO and land use factors (PERMANOVA, Table 7-8). The NMDS plot indicated, in particular, a divergent composition of the AMF community in the PC soil of the Lusignan LTO (which was the only arable soil included in this study; Fig. 2). This specific land use also featured the lowest (albeit non-significantly so) richness and Shannon index values (Table 9).

The intensification level did not have a significant effect when the four LTOs were analysed together (Table 7). However, the interaction between the LTO and intensity factors indicates some LTO-specific effect.

Factor	r^2	*p-value*
LTO	0.291	**0.000999**
season	0.02	0.3876
intensity	0.013	0.8412
land use	0.396	**0.000999**
LTO*season	0.049	0.234765
LTO*intensity	0.09030	**0.000999**
season*intensity	0.02	0.4396
season*land use	0.109	0.131868

Table 7. Results of PERMANOVA (1000 permutations) based on the Jaccard distance matrix of the AMF communities (combined dataset: spring+autumn). Significant results (p-value< 0.05) are in bold type.

	Lusignan *vs* Berchidda		Lusignan *vs* Veluwe		Lusignan *vs* Lancaster		Berchidda *vs* Veluwe		Berchidda *vs* Lancaster		Veluwe *vs* Lancaster	
Factor	r^2	*p-value*	r^2	*p-value*	r^2	*p-value*	r^2	*p-value*	r^2	*p-value*	r^2	*p-value*
LTO	0.304	**0.000999**	0.264	**0.000999**	0.221	**0.000999**	0.245	**0.000999**	0.16	**0.000999**	0.151	**0.000999**
season	0.025	0.973	0.038	0.6923	0.049	0.1708	0.02286	0.8831	0.051	0.08591	0.049	0.07892
intensity	0.045	0.6334	0.055	0.3227	0.052	0.1648	0.032	0.6963	0.038	0.2797	0.044	0.1389
land use	0.444	**0.000999**	0.425	**0.000999**	0.343	**0.000999**	0.345	**0.000999**	0.234	**0.000999**	0.248	**0.000999**
LTO*season	0.037	0.52048	0.028	0.737263	0.035	0.247752	0.035	0.404595	0.037	0.179820	0.038	0.117882
LTO*intensity	0.092	**0.022977**	0.101	**0.00999**	0.07	**0.004995**	0.066	**0.034965**	0.160	**0.000999**	0.055	**0.015984**
season*intensity	0.043	0.7213	0.034	0.7902	0.025	0.7812	0.044	0.4466	0.034	0.39960	0.031	0.39860
season*land use	0.134	0.215784	0.108	0.334665	0.099	0.119880	0.098	0.529471	0.089	0.312687	0.078	0.387612

Table 8. Results of PERMANOVA analyses (1000 permutations) on all LTO pairwise comparisons (Jaccard distance matrices of the combined dataset: spring+autumn). Significant results (p-value<0.05) are in bold type.

Group	*Richness*	*Shannon Index (H)*
PC	6.25 (4.35)	0.69 (0.39)
PG	18.75 (2.22)	2.02 (0.35)
IG	18.20 (2.59)	1.91 (0.42)
WP	20.00 (5.55)	2.18 (0.26)
I	20.00 (4.62)	2.05 (0.35)
U	19.50 (2.95)	2.15 (0.20)
High	20.57 (4.20)	2.00 (0.34)
Low	19.00 (3.79)	1.93 (0.43)
Lusignan	12.50 (7.41)	1.36 (0.79)
Berchidda	19.18 (4.35)	2.06 (0.35)
Lancaster	19.75 (3.78)	2.10 (0.28)
Veluwe	19.85 (3.93)	1.97 (0.37)
High intensity	17.69 (6.38)	1.80 (0.60)
Low intensity	19.38 (3.60)	2.08 (0.30)
Spring	17.54 (5.52)	1.95 (0.53)
Autumn	19.54 (4.75)	1.93 (0.46)

Table 9. Richness (mean number of taxa) and diversity (mean Shannon index) of soil AMF communities from each LTO, land use, intensity, season (combined dataset: spring+autumn). Standard deviation values are indicated within brackets. No statistically significant differences were found (non-parametric Kruskal-Wallis

Soil variables	r^2	*p-value*
Clay	0.128	**0.03**
Fine loam	0.2949	**0.001**
Coarse loam	0.3982	**0.001**
Fine Sand	0.337	**0.001**
Coarse sand	0.0802	0.121
Organic Carbon	0.2559	**0.002**
Total N	0.2131	**0.003**
C/N	0.1444	**0.022**
pH	0.4867	**0.001**
Phosphorus	0.1893	**0.012**
CEC cobaltiheamine	0.1095	0.057
Magnesium (Mg)	0.2267	**0.004**
Sodium (Na)	0.2254	**0.004**
Potassium (K)	0.0228	0.482
Iron (Fe)	0.0525	0.213
Manganese (Mn)	0.1218	**0.047**
Alluminium (Al)	0.201	**0.007**
Calcium (Ca)	0.0678	0.174

Table 10. Soil properties fitted onto NMDS ordination. R-squared and p-values were obtained from *envfit* analysis (999 permutations, vegan package). P-values< 0.05 are highlighted in bold type.

The *envfit* function of the vegan package (999 permutations) indicated a significant link of some environmental vectors (soil properties) to the ordination space of AMF community composition (Fig. 2, Table 10). The most significant ($p \leq 0.005$) soil characteristics were several soil textural characteristics, pH, organic C, total N, Mg and Na contents (Table 10).

Mantel tests carried out on the geographic distances and Jaccard dissimilarity matrices indicated that AMF community structures were also explained by geographic distances ($r_{spring+autumn\_data}$= 0.175, $p_{spring+autumn:data}$=0.005, r_{spring_data}= 0.3389, p_{spring_data}=0.001, r_{autumn_data}=0.3984, p_{autumn_data}=0.001).

2.4.2.2 *LTO-level analyses (local scale)*

Differences in AMF community composition (presence/absence data, Jaccard distances) were also analysed separately for each LTO.

For these analyses, LTO-specific subsamplings of the sequencing data were performed, yielding 850, 1550, 2800 and 2100 reads per sample for Lusignan, Berchidda, Lancaster and Veluwe respectively (Table S2 a-d). PERMANOVA analysis (Table 11) indicated a significant effect of the land use intensity on the AMF community composition in Lusignan, Berchidda and Veluwe sites. By contrast, land use intensity was not a significant factor for the Lancaster LTO, for which, instead, season was found to have a significant effect.

	Lusignan		*Berchidda*		*Lancaster*		*Veluwe*	
	r^2	*p*	r^2	*p*	r^2	*p*	r^2	*p*
Intensification level	0.29466	**0.01598**	0.17023	**0.04795**	0.06926	0.14785	0.15301	**0.03497**
Season	0.09469	0.59041	0.05266	0.86314	0.08907	**0.04595**	0.08377	0.41558
*Intensification level*season*	0.14777	0.28871	0.11410	0.28472	0.3319	0.85215	0.05769	0.709298

Table 11. PERMANOVA results (1000 permutations) for AMF soil community comparisons performed for each Long Term Observatory (Jaccard distance matrices of combined datasets: spring+autumn). Significant results (p-value<0.05) are in bold type

For each LTO the number of sequences belonging to each of the four Glomeromycota orders (based on Schüßler & Walker 2010, Redecker *et al.* 2013) is reported in Table 12 and Figure 3. At the three LTOs featuring a significant effect of the land use intensity (Lusignan, Berchidda and Veluwe), such effects involved all Glomeromycota orders. By contrast, at the Lancaster site, significant land use intensity effects were only found for Diversisporales and Archaeosporales (Table 12). Season also significantly affected all Glomeromycota orders at all sites (Table 12). Higher evenness of the four orders was found at Berchidda, Veluwe and Lancaster, whereas Lusignan featured a dominance of Glomerales (Fig. 3).

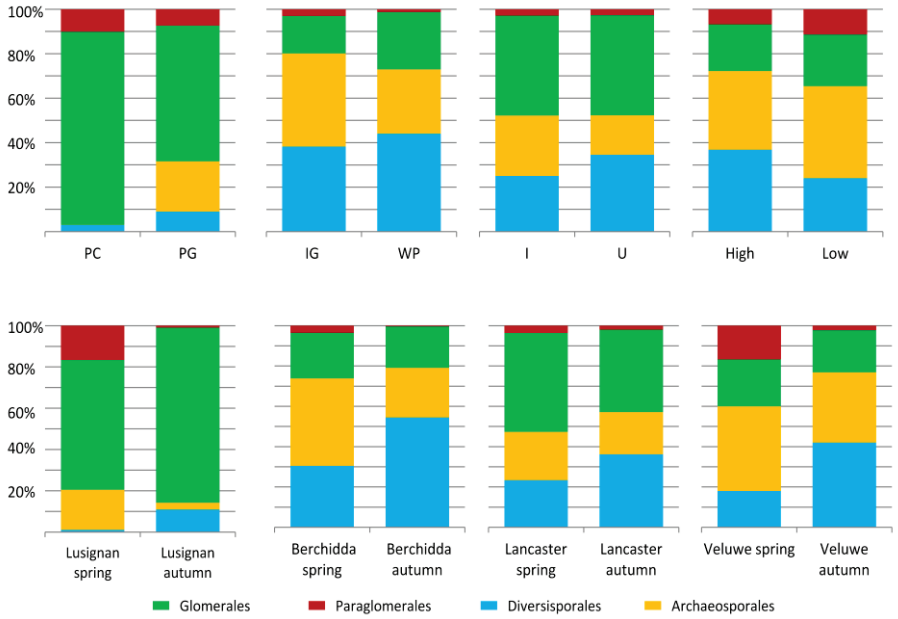

Figure 3. Percentages of total sequences assigned to a specific Glomeromycota order for each land use and season. Numbers of sequences and results of the statistical tests are reported in Table 12.

Land use intensity

		Diversisporales	Archaeosporales	Glomerales	Paraglomerales
Veluwe	High	5414*	5197*	3094*	995*
	Low	3024	5220	2923	1433
Lusignan	PC	103*	3*	2950*	344*
	PG	303	773	2072	252
Berchidda	IG	2966*	3244*	1302*	238*
	WP	4093	2691	2393	123
Lancaster	I	3128*	3395*	5619	358
	U	4319	2222	5621	338

Season

		Diversisporales	Archaeosporales	Glomerales	Paraglomerales
Veluwe	spring	2269*	5288*	2940*	2103*
	autumn	6169	5129	3077	325
Lusignan	spring	37*	662*	2137*	564*
	autumn	369	114	2885	32
Berchidda	spring	2838*	4028*	2102*	332*
	autumn	4221	1907	1593	29
Lancaster	spring	2923*	2993*	6133*	451*
	autumn	4524	2624	5107	245

Table 12. Order-level community composition of AMF assemblages. Data are numbers of sequences obtained for each of the four Glomeromycota orders (Diversisporales, Archaeosporales, Glomerales, Paraglomerales). For each LTO and order (column), significant differences between the two levels of intensification (land uses) and between the two sampling seasons were determined using chi-square tests (asterisk: $p < 0.05$). Prior to the analysis, replicates of each LTO had been rarefied at the same sequencing depth (Berchidda: 1550 sequences, Veluwe: 2100 sequences, Lancaster: 2800 sequences, Lusignan: 850 sequences).

The soil variables which best explained AMF community differences were forward-selected prior to variance partitioning analysis (Table S3). Such forward-selected soil variables varied depending on the LTO. No soil variable was selected for the Lancaster LTO. Overall, "soil" (soil physico-chemical properties), "intensity" (the land use intensification level), "season" (temporal variation) and their interactions accounted for 23.0% of total community variance. At Lusignan, Berchidda and Veluwe a greater proportion of variation in AMF communities was explained by soil properties (8.0-36.0%) and by the interaction between soil properties and land use intensity (5.0-26.0%) than by temporal variation (0.0-6.0%) and land use intensity *per se* (0.0-7.0%). In the case of Lancaster, by contrast, "season" explained a higher proportion of AMF community variation (3.4%) than either land use intensity (1.0%) or soil properties (0.0%). A large amount of AMF community structure variance could not be explained, indicating that other variables, which were not measured, were important drivers of AMF communities in the LTO soils under study.

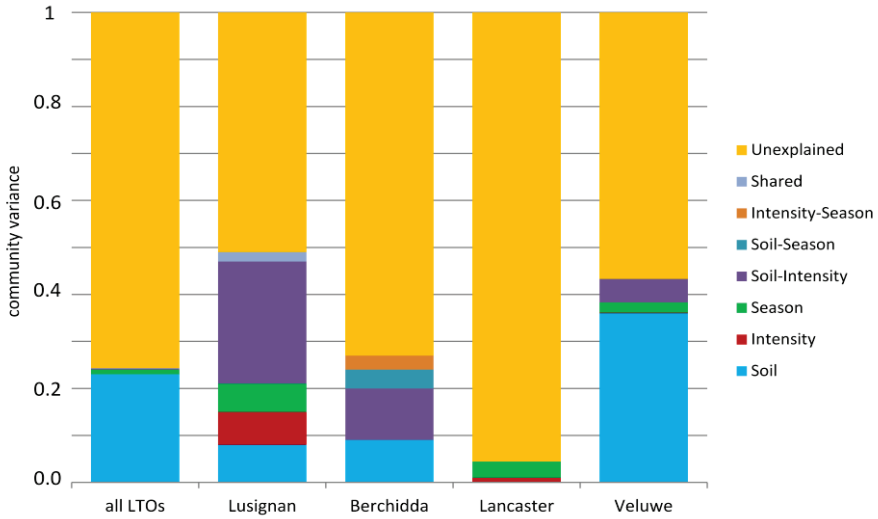

Figure 4. Partitioning of the variances of the AMF communities by explanatory variables. Barplots represent the variance explained by soil variables only (Soil), by land use intensity only (Intensity), by season only (Season), by soil & land use intensity & season (Shared), by two explanatory variables, and the unexplained variance (Unexplained).

To analyse possible correlations between soil parameters and AMF community composition at the LTO scale, Mantel tests were computed by using community distance matrices (Jaccard distances) and environmental distance matrices (Euclidean distances) obtained for each LTO singularly. No significant correlation was found for the Lusignan and Lancaster sites, whereas correlations were significant for Berchidda and Veluwe (Table 13).

LTO	*All properties*	
	r	*p*
Lusignan	0.0037	0.485
Berchidda	0.2253	**0.018**
Veluwe	0.4419	**0.003**
Lancaster	0.09863	0.183

Table 13. Mantel test results (999 permutations), bold value indicated significant correlation between two matrices ($p \leq 0.05$)

2.4.3 *Analyses at single species level (Indicator species analysis)*

Indicator species analysis was carried out to test whether single AMF taxa or combination of taxa could be found as being representative of a particular land use and/or group of land use and/or land use intensity. Although no effect of the season on the AMF communities (composition) was found (PERMANOVA analysis results), seasonal effects on single AMF taxa could not be ruled out. For this reason, we computed the indicator species analysis on the spring and autumn datasets separately. Table 14 reports the taxa / combinations of two taxa that yielded consistent results (IndVal.g$\geq$0.6, $p<0.05$) in both spring and autumn, as we hypothesize that a good indicator should be found associated to the same group of land uses in both spring and autumn.

Three single taxa and nine combinations of two taxa were found in both seasons as being associated to a specific land use type or to a group of land uses with a significant indicator value. No LTO-independent indicator of either low or high intensity was found.

			Spring				*Autumn*			
Land use group	*taxon1*	*taxon2*	A	B	IndVal.g	p-value	A	B	IndVal.g	p-value
PG	*Archaeosporaceae_sp*	*Septoglomus_viscosum*	1	1	1	0.005 **	1	1	1	0.014 *
	Ambispora_fennica	*Septoglomus_sp*	1	1	1	0.005 **	0.8333	1	0.913	0.016 *
	Rhizophagus_sp	*Septoglomus_sp*	0.8333	1	0.913	0.013 *	0.8333	1	0.913	0.016 *
	Archaeosporaceae_sp	*Septoglomus_sp*	0.8333	1	0.913	0.013 *	0.7143	1	0.845	0.025 *
	Claroideoglomus_claroideum	*Septoglomus_sp*	0.8333	1	0.913	0.013 *	0.7143	1	0.845	0.025 *
I+U	*Glomus_macrocarpum*		1	0.8	0.894	0.003 **	1	0.7	0.837	0.013 *
	Acaulospora_brasiliensis	*Glomus_macrocarpum*	1	0.8	0.894	0.003 **	1	0.7	0.837	0.013 *
	Archaeosporaceae_sp	*Glomus_macrocarpum*	1	0.8	0.894	0.003 **	1	0.7	0.837	0.013 *
	Glomus_macrocarpum	*Glomus_sp*	1	0.8	0.894	0.003 **	1	0.7	0.837	0.013 *
PC+PG	*Septoglomus_constrictum*		1	0.75	0.866	0.024 *	1	1	1	0.001 ***
High+I+IG+Low+ PG+U+WP	*Archaeosporaceae_sp*		1	1	1	0.009 **	1	0.9583	0.979	0.03 *
	Archaeosporaceae_sp	*Glomeraceae_sp3*	1	1	1	0.009 **	1	0.9583	0.979	0.03 *

Table 14. Results of Indicator Species Analysis. Specifity (A), sensitivity (B), Indicator species values (IndVal.g) and p-values are reported. Only taxa /combination of taxa consistent in the two seasons are reported. For PG, I, U, PC, High, Low, IG, WP see Table 1.

2.5 Discussion

In this study we investigated simultaneously AMF communities in the soils of a selection of four sites (Long Term Observatories, LTOs) across Europe. These European LTOs represent habitats with different types of soil, land use and intensification level at both the continental (among-LTO) and local (intra-LTO) scale. We identified soil AMF communities on the basis of 454 sequencing of ITS2 amplicons, which has the benefit of incorporating both the extraradical hyphae and spores in soil (Hempel *et al.* 2007, Dumbrell *et al.* 2010a). Our objectives were to explore the relative contributions of land use intensity, soil physico-chemical features and temporal (seasonal) variation to AMF diversity and community composition, and to identify broadly applicable indicators of land use intensification.

2.5.1 *Community-level patterns*

At the continental scale, the LTO and the land use type significantly influenced soil physico-chemical properties as well as AMF community composition, while season and land use intensification level *per se* (high- *vs* low- intensity) did not (PERMANOVA). Variance partitioning indicated that the greatest proportion (23%) of the variation in the AMF community composition explained by the variables under study was attributable to soil properties. In agreement with previous studies (Kivlin *et al.* 2011), we found a significant positive correlation between community dissimilarity and geographic distance ("distance decay relationship", Martiny *et al.* 2006). Spatial variation in species assemblages can be explained by either contemporary environmental conditions (present-day attributes of the environment) or historical contingencies (past events related to origin, dispersal and extinctions of species) (Martiny *et al.* 2006, Lindström *et al.* 2012). The relative influence of environmental versus historical factors seems to be related to the scale of sampling (Martiny *et al.* 2006), the legacy of historical separation overwhelming any effect of environmental factors at the intercontinental scale (on the order of tens of thousands of kilometres; e.g. Papke *et al.* 2003, Whitaker *et al.* 2003), whereas environmental effects have been repeatedly shown to significantly influence biotic composition at small spatial scales for which distance effects seem to be negligible (e.g. Horner-Devine *et al.* 2004, Kuske *et al.* 2002). It is instead at intermediate scales (10–3000 km), such as the one explored in our study, that the combined influence of both historical contingencies and contemporary ecological factors on microbial biogeography was most often detected (e.g. Green *et al.* 2004, Yannarell *et al.* 2005). AMF community composition in maize field soils in Zimbabwe was for example shown to be the product of both dispersal and environmental variables (Lekberg *et al.* 2007). Although we cannot rule out a role for dispersal limitation in the spatial patterning we observed, the results of the variance partitioning analysis suggest that

it may reflect the environmental heterogeneity of the study sites. Factoring in spatial autocorrelation at multiple scales in community variance partitioning (as in Horn *et al.* 2014) could help disentangling spatially from non-spatially structured effects of the environment on AMF communities.

AMF niche space is likely to be complex because of small-scale heterogeneity of soil (Veresoglou *et al.* 2013), and thus large-scale studies may overlook important drivers of local AMF community assembly (Horn *et al.* 2014). Furthermore, as previously mentioned, studies concentrated in a small area reduce confounding factors such as historical events and/or dispersal limitation, which are present in broad scale studies. At the local (intra-LTO) scale, we found significant effects of the intensification level (PERMANOVA) and soil characteristics (Mantel tests) on AMF communities in at least some of the four LTOs. At the Berchidda and Lusignan LTOs, where two types of land use are very different, the greater proportion of the explained variance in AMF communities was attributable to the interaction between soil properties and land use intensity. These findings indicate that the effects of land use intensification on AMF communities are mediated by corresponding change in soil parameters. In our study, it is possible that significant effects of the environment on AMF may be confounded with environmental effects on the host plants (as in Sharma *et al.* 2009). Berchidda and Lusignan were indeed the LTOs at which the two intensification levels were associated with the most dramatic changes in plant communities. Other members of the soil microbiota from the same LTOs have been shown to exhibit similar community patterns as we found here for AMF, for instance bacteria, archea and non-AMF fungi, indicating that that knowledge of land use practices *per se* is insufficient to predict the composition of soil microbial communities (Thomson *et al.* unpublished). Similarly, in northern China, Xiang *et al.* (2014) found that land use influence on AMF community composition was mediated by soil properties at landscape-scale (in particular, AMF communities from maize cultures exhibited less diverse communities than communities from grassland soils). Hazard and colleagues (2013), by using *Trifolium repens* and *Lolium perenne* bait plants to compare AMF communities in the soil of 40 geographically dispersed sites in Ireland representing different land uses and soil types, found that AMF community composition was influenced by abiotic variables (pH, rainfall and soil type), but not land use or geographical distance, suggesting that specific environmental variables of sites that vary within land uses have a stronger effect than land use itself on AMF communities. Local effects of land use intensity on AMF community composition where also observed in Estonia between intensive and sustainable arable lands (Moora *et al.* 2014).

Variance partitioning showed that the role of temporal factors in determining AMF community composition was instead marginal for three out of the four LTOs. In the case of Lancaster, by contrast, season (temporal variation) was the factor

accounting for the highest proportion of AMF community variation explained by the factors under study. The non-significant effects of temporal variation on AMF communities at Berchidda, Lusignan and Veluwe contrasted with the significant effects observed for specific AMF orders (Table 12). Given that these taxonomic groups may exhibit different dynamics, this could lead to diminished significance in the overall data set. Temporal shifts in AMF communities have been already described in grassland ecosystems (Dumbrell *et al.* 2011) as well in semiarid-arid agroecosystems (Bainard *et al.* 2014) and maple-associated communities (Helgason *et al.* 2014). By contrast, Santos-González *et al.* (2007) did not observe significant dynamics of AMF communities in a Swedish semi-natural grassland.

In our study, we certainly have missed out major environmental predictors of AMF communities. Indeed, most (51.0-95.6%) community-level variance could not be explained, indicating that other, unmeasured variables were important drivers of AMF communities. In AMF, phylogenetic community patterns can inform on assembly processes (HilleRisLambers *et al.* 2012, Roger *et al.* 2013, Horn *et al.* 2014, Moora *et al.* 2014) because AMF traits are phylogenetically conserved (Powell *et al.* 2009). Furthermore, environmentally-independent, stochastic events (such as population dynamics due to irregular, unpredictable environmental or demographic fluctuation) can deeply affect AMF assemblages (Lekberg *et al.* 2007, Dumbrell *et al*. 2010b, Verbruggen *et al*. 2012). For instance, Dumbrell *et al.* (2010b) suggested that chance-events could lead to a positive feedback mechanism on any taxon in the community, which could be random and self-reinforcing. Biotic interactions such as competition may also contribute to shaping community composition (Horn *et al.* 2014).

2.5.2 *Individual taxon-level patterns*

As previously mentioned, we also aimed at identifying individual land use features affecting individual AMF taxa across different soil types and climatic zones.

Similarly, Jansa and colleagues (2014) surveyed a broad range of Swiss agricultural soils with the aim of identifying AMF biomarkers of land use. In spite of the extensive coverage of large geographical and soil gradients, they did not identify any taxon suitable as a broadly applicable indicator. However, these authors profiled indigenous AMF communities in soils collected across Switzerland by means of quantitative realtime PCR with taxon-specific markers for six widespread AMF species, and acknowledge that rare species, which are likely to be affected by agricultural management practices (Verbruggen *et al.* 2012), were not quantified in their study, because they focused on a few, often dominant, AMF taxa (Maherali & Klironomos 2012). Furthermore, they pointed out that frequent tillage or massive overfertilization are not present in Swiss agriculture due to sanctions imposed by national legislation.

In spite of the different approach and high degree of intensification examined in our work, our study also failed to retrieve "universal" AMF indicators (individual taxa or combinations of taxa) of land use intensification. Similar results were obtained for bacteria, archea and non-AMF fungi in the same soil samples we analyzed. Taken together, the results obtained by our group as well as the other authors quoted above suggest that specific microbial indicators of change in biodiversity are likely to be dependent on local soil and climatic effects and the nature of the land use intensification (Thomson *et al.* unpublished).

In conclusion, the results obtained in our study suggest that alterations of soil features induced by land use intensification are the main drivers of changes in AMF communities in the soil. Several factors influence the structure of AMF assemblages and in general their effects are not "universal" but context-dependent. The responses of AMF taxa as well of AMF communities to a particular factor can indeed differ depending on the specific environmental framework.

2.6 Acknowledgments

DNA from the analysed soil samples was extracted by Pierre Plassart (INRA Dijon). We are grateful to Alberto Orgiazzi (EU-commission, JRC) for his useful suggestions. This work was supported by the European Commission project EcoFINDERS (FP7-64465).

2.7 References

Augé, R. (2001). Water relations, drought and vesicular-arbuscular mycorrhizal symbiosis. *Mycorrhiza* 11: 3–42.

Azcón-Aguilar, C., Barea, J. (1997). Arbuscular mycorrhizas and biological control of soil-borne plant pathogens–an overview of the mechanisms involved. *Mycorrhiza* 6:457–464.

An, G.H., Miyakawa, S., Kawahara, A., Osaki, M., Ezawa, T. (2008). Community structure of arbuscular mycorrhizal fungi associated with pioneer grass species *Miscanthus sinensis* in acid sulfate soils: habitat segregation along pH gradients. *Soil Science and Plant Nutrition* 54: 517–528.

Bai, C., He, X.L., Tang, H.L., Shan, B.Q., Zhao, L.L. (2009). Spatial distribution of AMF, glomalin and soil enzymes under the canopy of *Astragalus adsurgens* pall in the Mu US Sandland, China. *Soil Biology and Biochemistry* 41: 942–947.

Bainard, L.D., Koch, A.M., Gordon, A.M. *et al*. (2011) Influence of trees on the spatial structure of arbuscular mycorrhizal communities in a temperate tree-based intercropping system. *Agr Ecosyst Environ* 144:13–20.

Bainard, L. D., Bainard, J. D., Hamel, C., Gan, Y. (2014). Spatial and temporal structuring of arbuscular mycorrhizal communities is differentially influenced by abiotic factors and host crop in a semi-arid prairie agroecosystem. *FEMS Microbiology Ecology*, *88*(2), 333–44.

Berger, S. A., Krompass, D., Stamatakis, A. (2011). Performance, accuracy, and web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst. Biol.* 60:291–302.

Ciesielsky, H., Sterckeman, T. (1997). Determination of cation exchange capacity and exchangeable cations in soils by means of cobalt hexamine trichloride. Effects of experimental condictions. *Agronomie* 17:1–7.

Colwell, R.K. (2013). EstimateS: Statistical Estimation of species richness and shared species from samples. Version 9.0.0 User's Guide and application published at: http://purl.oclc.org/estimates

Dai M., Bainard, L.D., Hamel, C., Gan, Y., Lynch, D. (2013) Impact of land use on arbuscular mycorrhizal fungal communities in rural Canada. *Appl Environ Microbiol* 79: 6719–6729.

De Cáceres, M., Legendre, P. (2009). Associations between species and groups of sites: indices and statistical inference. *Ecology* 90:3566–3574.

Drigo, B., Pijl, A.S., Duyts, H., Kielak, A.M., Gamper, H.A., Houtekamer, M.J. (2010). Shifting carbon fl ow from roots into associated microbial communities in response to elevated atmospheric CO_2. *PNAS* 107(24):10938-10942

Dufrene, M., Legendre, P. (1997). Species Assemblages and Indicator Species: The Need for a flexible asymmetrical approach. *Ecol. Monogr*. 67:345.

Dumbrell, A.J., Nelson, M., Helgason, T., Dytham, C., Fitter, AH. (2010a). Relative roles of niche and neutral processes in structuring a soil microbial community. *ISME J*. 4:337–45.

Dumbrell, A.J., Nelson, M., Helgason, T., Dytham, C., Fitter, A.H. (2010b). Idiosyncrasy and overdominance in the structure of natural communities of arbuscular mycorrhizal fungi: is there a role for stochastic processes? *J. Ecol.* 98: 419–428.

Dumbrell, A.J., Ashton, P.D., Aziz, N. *et al.* (2011) Distinct seasonal assemblages of arbuscular mycorrhizal fungi revealed by massively parallel pyrosequencing. *New Phytol*. 190:794-804

Edgar,R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–1.

Fitter, A.H., Helgason, T., Hodge, A. (2011). Nutritional exchanges in the arbuscular mycorrhizal symbiosis: Implications for sustainable agriculture. *Fungal Biol. Rev.* 25:68–72.

Fitzsimons, M.S., Miller, R.M., Jastrow, J.D. (2008). Scale-dependent niche axes of arbuscular mycorrhizal fungi. *Oecologia* 158:117–27.

van der Gast, C.J., Gosling, P., Tiwari, B., Bending, G.D. (2011). Spatial scaling of arbuscular mycorrhizal fungal diversity is affected by farming practice. *Environ. Microbiol*. 13:241–9.

Green, J.L.; Holmes, A.J.; Westoby, M.; Oliver, I.; Briscoe, D.; Dangerfield, M.; *et al.* (2004). Spatial scaling of microbial eukaryote diversity. *Nature*, 432, 747–750.

Hazard, C., Gosling, P., van der Gast, C.J., Mitchell, D.T., Doohan, F.M., Bending, G.D. (2013). The role of local environment and geographical distance in determining community composition of arbuscular mycorrhizal fungi at the landscape scale. *ISME J.* 7:498–508.

van der Heijden, M. and Klironomos, J. (1998). Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* 74:69–72.

Hazard, C., Gosling, P., van der Gast, C. J., Mitchell, D. T., Doohan, F. M., Bending, G. D. (2013). The role of local environment and geographical distance in determining community composition of arbuscular mycorrhizal fungi at the landscape scale. *The ISME Journal* 7(3):498–508.

Helgason, T., Daniell, T.J., Husband, R., Fitter, A.H., Young, J.P. (1998). Ploughing up the wood-wide web? *Nature* 394:431.

Hempel, S., Renker, C., Buscot, F. (2007). Differences in the species composition of arbuscular mycorrhizal fungi in spore, root and soil communities in a grassland ecosystem. *Environ. Microbiol.* 9:1930–8.

Hildebrandt, U., Regvar, M., Bothe, H. (2007). Arbuscular mycorrhiza and heavy metal tolerance. *Phytochemistry* 68:139–46.

HilleRisLambers, J., Adler, P.B., Harpole, W.S., Levine, J.M., Mayfield, M.M. (2012). Rethinking community assembly through the lens of coexistence theory. *Annu Rev Ecol Evol Syst* 43: 227–248.

Horn, S., Caruso, T., Verbruggen, E., Rillig, M. C, Hempel, S. (2014). Arbuscular mycorrhizal fungal communities are phylogenetically clustered at small scales. *The ISME Journal* 8: 2231-2242.

Horner-Devine, M. C., Lage, M., Hughes, J. B. & Bohannan, B. J. M. (2004). A taxa–area relationship for bacteria. *Nature* 432: 750–753.

Jansa, J., Erb, A., Oberholzer, H.R., Smilauer, P., Egli, S. (2014). Soil and geography are more important determinants of indigenous arbuscular mycorrhizal communities than management practices in Swiss agricultural soils. *Molecular Ecology*, *23*(8), 2118–35

Kivlin, S.N., Hawkes, C.V., Treseder, K.K. (2011). Global diversity and distribution of arbuscular mycorrhizal fungi. *Soil Biol. Biochem*. 43:2294–2303.

Krüger, M., Stockinger, H., Krüger, C., Schüssler, A. (2009). DNA-based species level detection of Glomeromycota: one PCR primer set for all arbuscular mycorrhizal fungi. *New Phytol.* 183:212–23.

Kuske, C. R. *et al*. (2002).Comparison of soil bacterial communities in rhizospheres of three plant species and the interspaces in an arid grassland. *Appl. Environ. Microbiol.* 68: 1854–1863

Lekberg, Y., Koide, R.T. (2005). Is plant performance limited by abundance of arbuscular mycorrhizal fungi? A meta-analysis of studies published between 1988 and 2003. *New Phytol.* 168:189–204.

Lekberg ,Y., Koide, R.T., Rohr, J.R., Aldrich-Wolfe ,L,. Morton, J.B. (2007). Role of niche restrictions and dispersal in the composition of arbuscular mycorrhizal fungal communities. *J. Ecol.* 95:95–105.

Li, T., Hu, Y.J., Hao, Z.P., L,i H., Wang, Y.S., Chen, B.D. (2013). First cloning and characterization of two functional aquaporin genes from an arbuscular mycorrhizal fungus *Glomus intraradices. New Phytol*. 197:617–30.

Lin, X., Feng, Y., Zhang, H., Chen, R., Wang, J., Zhang, J., *et al.* (2012). Long-term balanced fertilization decreases arbuscular mycorrhizal fungal diversity in an arable soil in North China revealed by 454 pyrosequencing. *Environ. Sci. Technol*. 46:5764–71.

Lindström, E.S.; Langenheder, S. (2012). Local and regional factors influencing bacterial community assembly. *Environ. Microbiol. Rep.* 4: 1–9.

Liu, Y., Shi, G., Mao, L., Cheng, G., Jiang, S., Ma, X*., et al.* (2012). Direct and indirect influences of 8 yr of nitrogen and phosphorus fertilization on Glomeromycota in an alpine meadow ecosystem. *New Phytol*. 194:523–35.

Maherali, H., Klironomos, J.N. (2012) Phylogenetic and trait-based assembly of arbuscular mycorrhizal fungal communities. *PLoS ONE* 7, e36695.

Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., *et al.* (2006). Microbial biogeography: putting microorganisms on the map. *Nat. Rev. Microbiol.* 4:102–12.

Moebius-Clune, D.J., Anderson, Z.U., Pawlowska, T.E. (2013). Arbuscular mycorrhizal fungi associated with a single agronomic plant host across the landscape: The structure of an assemblage. *Soil Biol. Biochem*. 64:181–190.

Moora, M., Davison, J., Öpik, M., Metsis, M., Saks, U., Jairus,T., *et al.* (2014). Anthropogenic land use shapes the composition and phylogenetic structure of soil arbuscular mycorrhizal fungal communities. *FEMS Microbiol. Ecol.* 90:609–21.

Morris, E.K., Busco, F., Herbst, C., Meiners, T., Obermaier, E., Waschke, N.W.,Wubet, T., Rillig, M.C. (2013). Land use and host neighbor identity effects on arbuscular mycorrhizal fungal

community composition in focal plant rhizosphere. *Biodiversity and Conservation* 22:2193–2205.
Mummey, D.L., Rillig, M.C. (2006). The invasive plant species *Centaurea maculosa* alters arbuscular mycorrhizal fungal communities in the field. *Plant Soil* 288:81–90.
Oehl, F., Laczko, E., Bogenrieder, A., Stahr, K., Bösch, R., van der Heijden, M., *et al*. (2010). Soil type and land use intensity determine the composition of arbuscular mycorrhizal fungal communities. *Soil Biol. Biochem*. 42:724–738.
Oksanen, J., Blanchet, F., Kindt G.R, Legendre, P., *et al.* (2013). vegan: Community Ecology Package. http://cran.r-project.org/package=vegan.
Olsen, S.R. (1954). Esimation of available phosphorus in soils by extraction with sodium bicarbonate. *Circ. United States Dep. Agric*. 939:1–9.
Öpik, M., Moora, M., Liira, J., Zobel, M. (2006) Composition of root-colonizing arbuscular mycorrhizal fungal communities in different ecosystems around the globe. *Journal of Ecology* 94: 778–790.
Öpik, M., Vanatoa, A., Vanatoa, E. *et al.* (2010) The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). *New Phytologist*, 188: 223–241.
Öpik, M., Zobel, M., Cantero, J.J. *et al*. (2013) Global sampling of plant roots expands the described molecular diversity of arbuscular mycorrhizal fungi. *Mycorrhiza* 23, 411–430.
Orgiazzi, A., Lumini, E., Nilsson, R.H., Girlanda, M., Vizzini, A., Bonfante,P., Bianciotto, V. (2012). Unravelling soil fungal communities from different Mediterranean land-use backgrounds. *PLoS One* 7:e34847.
Papke, R.T., Ramsing, N.B., Bateson, M.M., Ward, D.M. (2003). Geographical isolation in hot spring cyanobacteria. *Environ. Microbiol.* 5:650–659.
Pereira e Silva, M.C.,Schloter-Hai, B., Schloter, M., van Elsas, J.D., Salles, J.F. (2013). Temporal dynamics of abundance and composition of nitrogen-fixing communities across agricultural soils. *PLoS One* 8:e74500.
Plassart ,P., Terrat, S., Thomson, B., Griffiths, R., Dequiedt, S., Lelievre,.M., *et al.* (2012). Evaluation of the ISO standard 11063 DNA extraction procedure for assessing soil microbial abundance and community structure. *PLoS One* 7:e44279.
Powell, J.R., Parrent, J.L., Hart, M.M., Klironomos, J.N, Rillig ,M.C., Maherali, H. (2009) Phylogenetic trait conservatism and the evolution of functional tradeoffs in arbuscular mycorrhizal fungi*. Proc Biol Sci* 276:237–4245.
Redecker, D. (2000). Glomalean fungi from the Ordovician. *Science* 289(80):1920–1921.
Redecker, D., Schüssler, A., Stockinger, H., Stürmer, S.L., Morton, J.B., Walker, C. (2013). An evidence-based consensus for the classification of arbuscular mycorrhizal fungi (Glomeromycota). *Mycorrhiza* 23:515–31.
Roger, A., Colard, A., Angelard, C., Sanders, I.R. (2013). Relatedness among arbuscular mycorrhizal fungi drives plant growth and intraspecific fungal coexistence. *ISME J* 7: 2137–2146.
Santos-Gonzalez, J.C., Finlay, R.D., Tehler, A. (2007). Seasonal dynamics of arbuscular mycorrhizal fungal communities in roots in a seminatural grassland. *Applied Environmental Microbiology* 73: 5613–5623.
Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., *et al.* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75:7537–41.
Schnoor, T.K., Lekberg, Y., Rosendahl, S., Olsson, P.A. (2011). Mechanical soil disturbance as a determinant of arbuscular mycorrhizal fungal communities in semi-natural grassland. *Mycorrhiza* 21:211–20.

Schüßler, A., Walker, C. (2010). The Glomeromycota: a species list with new families and new genera. Gloucester, UK.

Sharma, D., Kapoor, R., Bhatnagar, A.K. (2009). Differential growth response of *Curculigo orchioides* to native arbuscular mycorrhizal fungal (AMF) communities varying in number and fungal components. *Eur J Soil Biol* 45: 328–333.

Smith, S., Read, D. (2008). Mycorrhizal symbiosis. Academic Press.

Tedersoo, L., Bahram, M., Polme, S., Koljalg, U., Yorou, N. S., Wijesundera, R.,*et al*. (2014). Global diversity and geography of soil fungi. *Science*, 36 doi:10.1126/science.1256688

Turrini, A., Giovannetti, M. (2012). Arbuscular mycorrhizal fungi in national parks, nature reserves and protected areas worldwide: a strategic perspective for their in situ conservation. *Mycorrhiza* 22: 81–97.

Verbruggen, E., van der Heijden, M., Weedon, J.T, Kowalchuk, G., Röling, W.F.M.. (2012). Community assembly, species richness and nestedness of arbuscular mycorrhizal fungi in agricultural soils. *Mol. Ecol.* 21:2341–53.

Veresoglou, D.S., Caruso, T., Rillig, M.C. (2013). Modelling the environmental and soil factors that shape the niches of two common arbuscular mycorrhizal fungal families. *Plant Soil* 368: 507–518.

Whitaker, R. J., Grogan, D. W.,Taylor, J. W. (2003). Geographic barriers isolate endemic populations of hyperthermophilic Archaea. Science 301:976–978.

White, T., Brun, T., Lee, S., Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In:PCR Protocols: a guide to methods and applications., Innis, M., Gelfand, D., Sninsky, J., White, T., (eds), Academic Press: New York, pp. 315–322.

Wickham, H. (2009). ggplot2: elegant graphics for data analysis. Springer New York http://had.co.nz/ggplot2/book.

Xiang, D., Verbruggen, E., Hu, Y., Veresoglou, S.D., Rillig, M.C., Zhou, W., *et al.* (2014). Land use influences arbuscular mycorrhizal fungal communities in the farming-pastoral ecotone of northern China. *New Phytol.* 204:968–78.

Yannarell, A. C., Triplett, E. W. (2004).Within- and between-lake variability in the composition of bacterioplankton communities: investigations using multiple spatial scales. *Appl. Environ. Microbiol.* 70:214–223.

Zhang, L., Zhang, J., Christie, P., Li, X. (2009). Effect of inoculation with the arbuscular mycorrhizal fungus *Glomus intraradices* on the root-knot nematode *Meloidogyne incognita* in cucumber. *J. Plant Nutr* 32:967–979.

2.8 Supplementary materials

Table S1. List of the 57 Glomeromycota taxa identified in the analysed soil samples. Numbers of sequences obtained before subsampling for each land use in spring and autumn are reported.

		Lusignan				Berchidda				Lancaster				Veluwe			
Taxon	tot no. of seq.	PC spring	PC autumn	PG spring	PG autumn	IG spring	IG autumn	WP spring	WP autumn	I spring	I autumn	U spring	U autumn	High spring	High autumn	Low spring	Low autumn
Acaulospora alpina	8524	0	1	0	6	275	3718	146	2048	77	279	626	1291	1	55	1	0
Acaulospora brasiliensis	2458	0	0	0	8	0	8	2	156	240	184	821	1035	0	4	0	0
Acaulospora cavernata	19533	0	0	0	0	30	717	119	8163	210	2995	354	1027	457	2021	239	3201
Acaulospora laevis	62	0	0	0	0	0	0	0	0	0	0	62	0	0	0	0	0
Acaulospora sieverdingii	19072	0	0	0	17	685	5879	1005	5565	344	1584	443	1405	330	525	340	950
Acaulospora sp.	5396	0	0	0	1	16	553	116	3319	6	51	490	828	0	8	4	4
Ambispora appendicula	1353	0	4	48	9	69	69	16	0	2	52	274	668	74	68	0	0
Ambispora fennica	5948	592	10	275	123	819	297	372	1430	91	0	249	333	901	165	259	32
Ambispora sp.	3671	0	0	11	1	9	1	2	0	158	1135	8	88	48	517	820	873
Archaeospora sp.	1916	1	0	0	0	13	15	165	34	6	489	5	371	53	176	254	334
Archaeospora trappei	599	0	3	77	4	167	136	42	90	3	0	25	20	21	2	1	8
Archaeosporaceae sp.	81740	0	1	890	459	2282	3176	1714	12575	3626	23021	2032	6500	2522	10080	4167	8695
Archaeosporales	1706	0	0	0	0	3	22	0	3	43	1166	48	338	1	10	8	64
Claroideoglomeraceae sp.	49	0	0	0	0	0	0	8	3	0	7	7	17	1	4	0	2
Claroideoglomus claroideum	12599	0	2	162	957	111	106	21	587	2325	4866	718	1496	401	645	111	91
Claroideoglomus etunicatum	589	0	0	0	2	146	1	29	81	1	167	0	1	10	84	0	67
Claroideoglomus sp.1	196	0	0	0	0	0	0	0	0	124	67	0	3	0	0	0	2

Claroideoglomus sp.2	5174	0	0	70	390	71	19	55	1008	380	2300	159	440	16	209	0	57
Claroideoglomus sp.3	10973	0	0	4	113	302	40	71	696	222	3723	109	4621	13	165	11	883
Claroideoglomus sp.4	761	0	0	8	101	0	3	7	0	50	492	46	31	1	17	0	5
Claroideoglomus sp.5	3107	0	0	0	0	0	1	0	0	29	520	91	1006	225	492	380	363
Claroideoglomus sp. W3349	7	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0
Diversispora celata	42	0	0	1	41	0	0	0	0	0	0	0	0	0	0	0	0
Diversispora eburnea	2727	0	0	41	4	0	0	4	9	32	1709	1	128	34	687	3	75
Diversispora epigaea	31859	0	400	57	1012	74	124	188	2342	280	14896	378	4954	488	5695	56	915
Diversispora sp.	7	0	0	0	0	0	0	0	0	0	0	0	2	0	5	0	0
Diversispora sp. W5257	2777	0	0	1	140	0	0	2	9	8	298	4	220	290	991	8	806
Funneliformis caledonius	8732	969	5823	58	366	157	1	99	0	123	1080	19	17	3	15	0	2
Funneliformis geosporus	18	0	0	0	10	2	0	0	0	0	6	0	0	0	0	0	0
Funneliformis mosseae	5826	0	1056	500	3531	0	0	0	1	1	8	151	0	493	74	5	6
Funneliformis sp.	155	0	0	0	0	0	0	5	21	0	73	0	37	0	0	14	5
Gigaspora sp.1	1340	0	0	0	0	0	1163	141	0	0	19	0	17	0	0	0	0
Gigaspora sp.2	5060	0	0	0	0	1	8	0	0	404	2233	23	1119	15	224	188	845
Gigasporaceae sp.	1405	0	0	0	0	0	0	54	954	1	0	1	3	266	63	29	34
Glomeraceae sp.1	1809	0	0	0	0	26	12	9	98	21	786	1	665	0	2	0	189
Glomeraceae sp.2	97	0	0	0	14	2	0	0	0	0	45	0	13	2	16	0	5
Glomeraceae sp.3	21637	75	2	93	247	315	591	329	6783	526	4044	272	764	1703	1358	609	3926
Glomeraceae sp.4	4750	456	7	63	111	43	141	250	780	344	276	653	350	338	160	92	686

Glomeraceae sp.5	267	0	0	0	0	0	0	0	0	1	0	0	3	0	65	67	131
Glomus cerebriforme	149	0	0	29	97	0	0	0	0	0	6	0	12	1	0	4	0
Glomus macrocarpum	825	0	0	0	0	0	0	0	0	26	265	50	482	0	2	0	0
Glomus sp.	28965	0	0	1	18	325	350	296	5605	1350	8978	3073	7212	92	209	165	1291
Glomus sp. W3347	432	0	0	0	0	0	0	0	0	278	140	0	0	0	1	9	4
Pacispora scintillans	24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	20
Paraglomus brasilianum	813	0	0	0	0	0	0	0	0	0	0	2	0	86	6	280	439
Paraglomus laccatum	8583	364	9	297	212	308	83	141	123	431	2366	525	461	1380	106	1558	219
Paraglomus sp.	83	0	0	0	0	0	0	0	0	1	45	1	14	21	0	1	0
Redeckera sp.1	5	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0
Redeckera sp.2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0
Rhizophagus irregularis	378	0	25	32	16	0	0	26	27	18	0	27	0	165	8	10	24
Rhizophagus sp.	2157	26	12	160	234	28	62	184	480	55	9	327	178	132	83	28	159
Scutellospora spinosissima	64	0	0	0	0	2	0	0	4	0	1	1	50	0	0	1	5
Cetraspora gilmorei	3389	0	0	0	171	729	33	116	1893	0	0	447	0	0	0	0	0
Scutellospora sp.	3037	0	0	0	1	192	208	149	344	93	170	92	286	320	191	521	470
Septoglomus constrictum	1208	74	173	145	815	0	0	0	0	0	0	0	0	0	0	0	1
Septoglomus sp.	492	0	13	8	39	0	0	0	0	0	161	41	229	0	0	0	1
Septoglomus viscosum	185	0	25	23	136	0	0	0	1	0	0	0	0	0	0	0	0
total	**324,732**	**2557**	**7566**	**3054**	**9406**	**7202**	**17.537**	**5883**	**55,232**	**11.930**	**80.712**	**12.656**	**38.742**	**10.904**	**25.215**	**10.247**	**25.889**

Table S2. List of the 57 Glomeromycota taxa identified in the analysed soil samples at each LTO. For each taxa is reported the number of sequences obtained for each land use in spring and autumn after LTO-specific subsampling. Numbers of sequences obtained after subsampling at even sequencing depth for each land use (Lusignan: 850 sequences/sample; Berchidda: 1550 sequences/sample; Lancaster: 2800 sequences/sample; Veluwe: 2100 sequences/sample) in spring and autumn are reported.

a) Lusignan LTO, PC: Permanent Culture, PG: Permanent Grassland

	PC2_spring	PC3_spring	PG2_spring	PG3_spring	PC1_autumn	PC3_autumn	PG1_autumn	PG4_autumn
Acaulospora_sieverdingii	0	0	0	0	0	0	0	3
Ambispora_appendicula	0	0	22	26	1	0	0	1
Ambispora_fennica	0	0	109	128	1	0	6	14
Ambispora_sp	0	0	6	5	0	0	0	0
Archaeospora_trappei	0	0	0	58	0	1	0	0
Archaeosporaceae_sp	0	0	59	249	0	0	55	35
Claroideoglomus_claroideum	0	0	14	15	0	0	235	14
Claroideoglomus_etunicatum	0	0	0	0	0	0	1	0
Claroideoglomus_sp2	0	0	41	0	0	0	91	1
Claroideoglomus_sp3	0	0	4	0	0	0	34	0
Claroideoglomus_sp4	0	0	0	0	0	0	15	0
Diversispora_celata	0	0	0	0	0	0	0	7
Diversispora_eburnea	0	0	0	3	0	0	0	1
Diversispora_epigaea	0	0	0	33	11	92	148	55
Diversispora_sp_W5257	0	0	0	1	0	0	6	18
Funneliformis_caledonius	0	793	0	0	748	509	93	0
Funneliformis_geosporus	0	0	0	0	0	0	1	0
Funneliformis_mosseae	0	0	364	86	82	184	0	508
Glomeraceae_sp2	0	0	0	0	0	0	0	3

Glomeraceae_sp3	73	0	10	25	0	1	27	19
Glomeraceae_sp4	436	0	7	14	0	0	12	4
Glomus_cerebriforme	0	0	3	0	0	0	22	0
Paraglomus_laccatum	341	0	105	118	0	3	18	11
Rhizophagus_irregularis	0	0	14	17	5	2	0	5
Rhizophagus_sp	0	0	15	27	0	2	27	4
Cetraspora_gilmorei	0	0	0	0	0	0	0	27
Scutellospora_sp	0	0	0	0	0	0	0	1
Septoglomus_constrictum	0	57	70	38	2	49	42	108
Septoglomus_sp	0	0	3	1	0	4	2	3
Septoglomus_viscosum	0	0	4	6	0	3	15	8

b) Berchidda LTO, IG: Intensive Grassland, WP: Wooded Pasture

	IG1_spring	IG2_spring	IG3_spring	WP1_spring	WP2_spring	WP3_spring	IG2_autumn	IG3_autumn	WP1_autumn	WP2_autumn	WP3_autumn
Acaulospora_alpina	0	0	136	15	6	124	0	481	166	0	0
Acaulospora_brasiliensis	0	0	0	1	0	1	0	1	18	0	0
Acaulospora_cavernata	25	0	0	3	5	110	112	42	173	333	0
Acaulospora_sieverdingii	0	142	226	451	189	133	229	655	309	65	0
Acaulospora_sp	0	12	1	36	22	35	1	50	105	4	481
Ambispora_appendicula	0	0	32	12	0	0	0	4	0	0	0
Ambispora_fennica	44	368	99	96	144	0	83	4	17	78	0
Ambispora_sp	1	0	7	2	0	0	0	0	0	0	0
Archaeospora_sp	12	0	0	2	106	0	5	1	6	0	0
Archaeospora_trappei	101	26	8	10	20	2	29	2	1	5	0
Archaeosporaceae_sp	500	837	253	431	558	354	723	96	514	333	0
Archaeosporales	1	2	0	0	0	0	6	0	0	0	0

Claroideoglomeraceae_sp	0	0	0	0	3	3	0	0	0	0	1
Claroideoglomus_claroideum	4	0	45	3	4	6	22	2	1	25	2
Claroideoglomus_etunicatum	101	22	0	11	5	0	0	0	5	0	0
Claroideoglomus_sp2	19	5	24	4	33	9	2	2	4	0	238
Claroideoglomus_sp3	46	2	123	13	36	1	8	2	25	0	72
Claroideoglomus_sp4	0	0	0	0	7	0	0	1	0	0	0
Claroideoglomus_sp5	0	0	0	0	0	0	1	0	0	0	0
Diversispora_eburnea	0	0	0	0	2	2	0	0	1	0	0
Diversispora_epigaea	7	1	27	46	46	61	44	0	26	83	60
Diversispora_sp_W5257	0	0	0	0	0	1	0	0	2	0	0
Funneliformis_caledonius	146	0	0	25	40	0	0	0	0	0	0
Funneliformis_geosporus	2	0	0	0	0	0	0	0	0	0	0
Funneliformis_mosseae	0	0	0	0	0	0	0	0	0	0	0
Funneliformis_sp	0	0	0	3	0	1	0	0	6	0	0
Gigaspora_sp1	0	0	0	0	17	107	0	142	0	0	0
Gigaspora_sp2	0	0	1	0	0	0	2	1	0	0	0
Gigasporaceae_sp	0	0	0	2	39	0	0	0	0	51	0
Glomeraceae_sp1	6	0	8	2	1	6	1	1	9	0	7
Glomeraceae_sp2	1	0	0	0	0	0	0	0	0	0	0
Glomeraceae_sp3	103	41	59	59	86	135	131	26	53	282	130
Glomeraceae_sp4	17	6	10	47	39	126	34	3	15	26	23
Glomus_sp	4	23	133	45	41	168	74	15	73	220	0
Paraglomus_laccatum	142	49	30	79	16	16	17	0	4	5	3
Rhizophagus_irregularis	0	0	0	2	6	12	0	0	2	2	0
Rhizophagus_sp	6	10	3	12	49	89	2	5	7	33	0

Scutellopsora_spinosissima	0	1	0	0	0	0	0	0	0	0	0
Cetraspora_gilmorei	211	0	260	93	0	0	0	3	1	0	473
Scutellospora_sp	51	3	65	45	30	48	24	11	7	5	60

c) Veluwe, High: short term abandoned grassland. Low: long term abandoned grassland

	High1_spring	High2_spring	High3_spring	Low1_spring	Low2_spring	Low3_spring	High1_autumn	High2_autumn	High3_autumn	High4_autumn	Low1_autumn	Low2_autumn	Low3_autumn
Acaulospora alpina	0	0	0	1	0	0	0	0	38	0	0	0	0
Acaulospora brasiliensis	0	0	0	0	0	0	1	0	1	0	0	0	0
Acaulospora cavernata	176	0	101	126	0	46	208	0	806	10	36	682	0
Acaulospora sieverdingii	135	0	57	175	0	70	131	0	0	0	7	188	0
Acaulospora sp.	0	0	0	0	2	0	0	0	0	0	0	0	1
Ambispora appendicula	8	33	0	0	0	0	8	9	7	0	0	0	0
Ambispora fennica	119	314	23	191	0	0	17	42	2	0	2	3	0
Ambispora sp.	6	16	5	21	409	22	15	26	2	89	1	10	140
Archaeospora sp.	1	22	2	6	141	1	14	6	13	23	1	12	48
Archaeospora trappei	2	7	2	0	0	0	2	0	0	0	5	0	0
Archaeosporaceae sp.	461	571	448	420	1098	934	726	821	132	1199	154	647	938
Archaeosporales	0	0	1	0	4	0	2	0	1	0	0	0	12
Claroideoglomeraceae sp.	1	0	0	0	0	0	1	0	0	1	0	0	0
Claroideoglomus claroideum	105	2	170	91	0	0	49	162	17	29	4	14	0
Claroideoglomus etunicatum	0	0	10	0	0	0	22	0	0	0	0	15	0
Claroideoglomus sp.2	0	0	14	0	0	0	1	0	6	44	11	2	5
Claroideoglomus sp.3	0	0	13	8	0	0	4	0	9	28	98	0	133
Claroideoglomus sp.4	0	0	0	0	0	0	0	0	7	1	1	1	0

Claroideoglomus sp.5	1	83	56	254	5	24	91	7	11	3	1	90	3
Diversispora eburnea	0	4	25	3	0	0	55	40	62	77	24	4	0
Diversispora epigaea	24	126	176	38	0	2	384	784	498	382	359	40	1
Diversispora sp.	0	0	0	0	0	0	2	0	0	1	0	0	0
Diversispora sp. W5257	5	2	264	3	0	4	16	33	327	81	377	11	0
Funneliformis caledonius	0	2	0	0	0	0	0	0	10	0	0	0	0
Funneliformis mosseae	0	164	145	0	0	1	16	3	0	1	0	2	0
Funneliformis sp.	0	0	0	0	9	0	0	0	0	0	0	0	2
Gigaspora sp.2	7	0	0	41	5	73	48	0	0	0	147	95	19
Gigasporaceae sp.	115	4	7	11	5	0	11	5	0	0	2	2	2
Glomeraceae sp.1	0	0	0	0	0	0	0	0	0	0	18	0	28
Glomeraceae sp.2	0	1	1	0	0	0	6	2	0	0	0	2	0
Glomeraceae sp.3	580	207	125	152	118	105	167	120	37	78	334	172	405
Glomeraceae sp.4	105	38	28	36	6	10	25	17	13	5	178	27	52
Glomeraceae sp.5	0	0	0	5	35	0	4	0	0	13	0	0	31
Glomus cerebriforme	0	0	1	0	0	3	0	0	0	0	0	0	0
Glomus macrocarpum	0	0	0	0	0	0	0	0	1	0	0	0	0
Glomus sp.	41	3	2	3	68	18	23	1	37	24	26	11	209
Glomus sp. W3347	0	0	0	2	0	0	1	0	0	0	0	1	0
Pacispora scintillans	0	0	0	3	0	0	0	0	0	0	12	0	0
Paraglomus brasilianum	2	0	84	201	0	0	0	0	3	0	159	0	30
Paraglomus laccatum	33	479	318	230	20	714	1	10	34	10	50	22	6
Paraglomus sp.	2	0	19	1	0	0	0	0	0	0	0	0	0
Rhizophagus irregularis	0	1	1	7	0	0	0	2	1	0	11	0	1

Rhizophagus sp.	51	7	0	17	2	3	6	4	25	1	32	16	5
Scutellopsora spinosissima	0	0	0	0	0	1	0	0	0	0	3	0	0
Scutellospora sp.	120	14	2	54	173	69	43	6	0	0	47	31	29

d) Lancaster, I: Improved grassland, U: Unimproved grassland

	I1_sprin g	I2_sprin g	I3_sprin g	I4_sprin g	I5_sprin g	U1_sprin g	U2_sprin g	U3_sprin g	U4_sprin g	U5_sprin g	I1_autumn	I2_autumn	I3_autumn	I4_autumn	I5_autumn	U1_autumn	U2_autumn	U3_autumn	U4_autumn	U5_autumn
Acaulospora_alpina	0	33	2	0	0	218	99	70	0	27	4	0	9	6	7	3	36	225	0	3
Acaulospora_brasiliensis	0	14	2	0	95	118	119	30	87	92	1	0	7	0	0	17	15	73	71	3
Acaulospora_cavernata	0	87	0	2	0	301	0	16	0	0	0	154	169	40	0	96	0	62	1	0
Acaulospora_laevis	0	0	0	0	0	1	27	0	0	0	0	0	0	0	0	0	0	0	0	0
Acaulospora_sieverdingii	0	0	0	0	168	0	0	0	120	34	0	1	60	0	24	136	0	67	0	0
Acaulospora_sp	0	1	0	0	2	1	116	126	0	0	0	0	5	0	0	0	60	57	0	0
Ambispora_appendicula	0	0	1	0	0	0	141	0	0	0	0	0	4	0	0	76	0	7	0	0
Ambispora_fennica	0	1	0	0	46	6	141	0	0	0	0	0	0	0	0	42	0	18	1	0
Ambispora_sp	17	50	5	3	2	0	1	0	2	0	61	2	3	5	0	2	2	0	5	0
Archaeospora_sp	0	2	1	0	0	0	0	1	0	2	5	3	15	0	1	0	16	0	29	0
Archaeospora_trappei	0	0	0	1	1	0	0	1	4	7	0	0	0	0	0	0	0	0	3	0
Archaeosporaceae_sp	406	707	72	199	367	5	79	18	490	178	585	212	306	152	74	96	359	57	342	32
Archaeosporales	11	1	0	0	10	0	0	0	6	8	30	4	20	10	0	3	32	0	10	0
Claroideoglomeraceae_sp	0	0	0	0	0	2	0	0	2	0	0	2	0	0	0	0	1	6	0	0
Claroideoglomus_claroideum	282	12	701	137	121	1	2	271	12	23	1	53	67	96	202	82	23	91	0	75
Claroideoglomus_etunicatum	0	0	0	0	0	0	0	0	0	0	0	1	1	0	5	0	0	1	0	0

Claroideoglomus_sp1	0	49	4	0	0	0	0	0	0	0	0	12	0	0	2	0	1	0	0	0
Claroideoglomus_sp2	109	0	39	34	10	0	15	47	6	2	1	8	23	51	103	1	4	12	29	39
Claroideoglomus_sp3	13	4	20	82	25	0	49	4	0	0	1	123	132	91	82	52	180	7	264	181
Claroideoglomus_sp4	12	0	5	16	0	0	0	21	1	0	0	0	10	0	23	0	3	0	0	3
Claroideoglomus_sp5	0	9	1	0	0	0	0	0	0	43	9	12	14	0	0	0	30	2	122	0
Claroideoglomus_sp_W3349	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0
Diversispora_eburnea	0	0	0	30	0	0	0	0	1	0	0	21	16	54	75	0	4	12	0	2
Diversispora_epigaea	51	0	8	150	1	92	0	28	13	83	209	288	244	338	246	129	266	193	0	288
Diversispora_sp_W5257	1	0	0	5	0	0	2	0	0	0	1	2	4	5	11	0	3	14	0	35
Funneliformis_caledonius	0	0	0	116	0	0	0	0	7	0	0	0	39	149	0	3	0	0	0	0
Funneliformis_mosseae	0	0	0	0	1	1	2	0	50	0	0	0	0	1	1	0	0	0	0	0
Funneliformis_sp	0	0	0	0	0	0	0	0	0	0	1	5	0	5	0	1	2	0	0	0
Gigaspora_sp1	0	0	0	0	0	0	0	0	0	0	1	6	0	0	0	0	3	0	0	0
Gigaspora_sp2	0	0	223	0	3	0	0	0	6	0	0	47	0	0	134	0	0	0	1	351
Gigasporaceae_sp	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Glomeraceae_sp1	1	1	0	4	0	0	1	0	0	0	0	30	43	0	9	13	24	0	14	71
Glomeraceae_sp2	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	2	0	0	0
Glomeraceae_sp3	36	49	20	108	77	20	27	23	18	70	83	52	30	95	45	20	25	21	49	15
Glomeraceae_sp4	57	8	9	15	86	91	6	173	34	29	0	2	2	6	13	11	3	24	13	4
Glomeraceae_sp5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Glomus_cerebriforme	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	2	0	0
Glomus_macrocarpum	0	1	7	0	4	10	6	7	1	8	0	0	3	0	9	40	11	26	3	0

Glomus_sp	0	198	112	330	160	368	411	182	182	499	198	124	13	144	140	382	104	252	277	35
Glomus_sp_W3347	109	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Paraglomus_brasilianum	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0
Paraglomus_laccatum	137	17	5	2	22	1	0	128	1	135	57	80	9	2	25	1	35	1	15	17
Paraglomus_sp	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	0	0	0	1	0
Rhizophagus_irregularis	5	0	3	0	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0
Rhizophagus_sp	3	6	0	11	13	14	5	75	39	8	0	0	1	0	0	7	6	18	0	1
Cetraspora_gilmorei	0	0	0	0	0	0	0	0	131	0	0	0	0	0	0	0	0	0	0	0
Scutellospora_sp	0	0	9	5	35	0	0	0	37	0	0	2	0	0	9	3	0	0	0	75
Scutellospora_spinosissima	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
Septoglomus_sp	0	0	0	0	0	0	0	16	0	0	0	1	0	0	10	34	0	0	0	0

all LTOs	Coarse loam content ,C:N ratio, total Nitrogen content, Iron, Potassium
Lusignan	Phosphorus , Coarse sand content
Berchidda	Phosphorus , Organic Carbon, Alluminium
Veluwe	Coarse sand content, pH, Potassium, C:N ratio, Alluminium
Lancaster	No variables selected

Table S3. Forward-selected environmental variables used to represent soil properties in variance partitioning analyses.

Chapter III

How will soil fungi actively decomposing plant litter respond to predicted rainfall reduction? The case of a Mediterranean forest

3.1 Foreword

Besides changes in land uses, forecasted global or local climate changes are expected to strongly impact terrestrial ecosystems functioning. In this Chapter we focused on the impact of changes in precipitation input in a Mediterranean forest ecosystem. Mediterranean areas are indeed predicted to experience in the future significant reduction in rainfall.

Unlike in the previous Chapter, the present study targets the whole soil fungal community (not a specific phylum or taxonomic group) and a specific subset of it that participates to plant organic matter degradation (that represents a specific, but polyphyletic, fungal functional group). In order to investigate if and how fungi actively decomposing the most abundant plant polymers (lignin, cellulose and hemicellulose) are affected by reduction in precipitation, we implemented a targeted metatranscriptomics approach. This environmental genomics approach uses soil RNA as starting material, instead of soil DNA (metagenomics approach, used in *Chapter II*), thus allowing the characterization of biologically active organisms. This approach is based on the metabarcoding of several gene families encoding fungal enzymes active on different plant cell wall polymers. It was therefore necessary to design and evaluate the performance of different gene family-specific PCR primers. I participated to this technical development presented in *Chapter V* focusing on three gene families.

Title: How will soil fungi actively decomposing plant litter respond to predicted rainfall reduction? The case of a Mediterranean forest

Claudia Bragalini[1,2], Patricia Luis[2], Elsa Prudent[2], Richard Joffre[3], Roland Marmeisse[1,2*], Mariangela Girlanda[1,4*]

Affilliations:
[1]*Department of Life Sciences and Systems Biology, University of Turin,*
[2]*Université Claude Bernard Lyon1, UMR CNRS-5557, USC INRA-1364,*
[3]*Centre d'Ecologie Fonctionnelle et Evolutive, CEFE CNRS, Montpellier*
[4]*CNR-Istituto per la Protezione Sostenibile delle Piante, UOS Turin (CNR-IPSP), Italy*
*These authors contributed equally to the work

Manuscript to be submitted

3.2 Introduction

Establishing the connection between microbial diversity and community structure and function has been a longstanding yet elusive goal in microbial and soil ecology. Making such a connection is especially critical for predicting how communities and functions will respond to environmental changes (Allison & Martiny 2008, Trivedi *et al.* 2013). Observations from diverse forest soils indicate that environmental factors, such as temperature, water availability and substrate quality, represent important drivers of soil microbial community composition and activity (Aponte *et al.* 2011, Kaiser *et al.* 2010, Landesman & Dighton 2011).

Degradation of plant polymers is a key microbial function and step in the global terrestrial carbon cycle that channels plant litter into microbial biomass, where it can be mineralized to CO_2 or stabilized as soil carbon (Kandeler *et al.* 2005). Soil organic matter (SOM) degradation is also essential for the recycling of essential nutrients (e.g. organic forms of N and P) trapped and complexed in dead plant biomass. In forest ecosystems saprotrophic fungi, but also to some extent symbiotic ectomycorrhizal ones (Bödeker *et al.* 2009, Rineau *et al.* 2013, Talbot *et al.* 2013, Lindahl & Tunlid 2014), are largely responsible for the breakdown (either hydrolysis or chemical modification) of the most abundant plant polymers cellulose, hemicelluloses and lignin (Berg & McClaugherty, 2008). For instance Schneider *et al.* (2012), using a metaproteomics approach, revealed that fungi were the main producers of extracellular hydrolytic enzymes in a beech (*Fagus sylvatica*) litter.

In the recent years, a number of studies have addressed the impact of global changes on soil biota and the provided ecosystem functions/services (García-Palacios *et al.* 2014). However, few of these studies focused on the impact of altered

precipitation regimes on litter decomposition (but see García-Palacios *et al.* 2014). Global climate change is indeed predicted to significantly alter precipitation and drought patterns, resulting in more extreme conditions (IPCC 2007). This should especially impact Mediterranean ecosystems (which are recognized as biodiversity hotspots). In the Mediterranean basin rainfall is indeed expected to decline by up to 30% in the warm season (Giorgi & Lionello 2008) and become more irregular by the end of 21st century (Gibelin & Déqué 2003, IPCC 2007, Giorgi & Lionello 2008), which will add up to the long drought periods that this area naturally experiences in summer.

Different studies have reported that both the structures and functionality of microbial communities, including fungal ones, are highly responsive to changes in soil moisture within and across years (Toberman *et al.* 2008, Bell *et al.* 2009, Baldrian *et al.* 2010, Castro *et al.* 2010, Hawkes *et al.* 2011, Cregger *et al.* 2012). Zeglin *et al.* (2013) showed that altered precipitation regimes in a native tallgrass prairie affects the function and the composition of soil microbial communities at different time scales, from days (rainfall event) to months (seasonal drought) to years (global alteration of the expected precipitation regime). Hawkes *et al.* (2011) concluded that fungal communities during low soil moisture periods were more diverse and abundant than during high soil moisture ones. Furthermore, these responses in fungal communities were rapid, reversible, and repeatable, thus highlighting the community plasticity in terms of environmental or physiological tolerances (Cruz-Martinez *et al.* 2009). Fundamentally, soil moisture is a pivotal driver of fungal biomass (Frey *et al.* 1999) and the communities respond rapidly to changes in available moisture (Baldrian *et al.* 2013).

Studies examining shifts in microbial communities commonly employ molecular surveys, often by targeting phylogenetically informative genes, such as the ribosomal (rRNA) ones. While these markers can provide insights into how biodiversity and community composition respond to environmental changes, they may not accurately reflect functional shifts, that may result in changes in ecosystem functions such as plant organic matter (POM) decomposition. Indeed, as ligninocellulolytic microorganisms represent a heterogeneous and polyphyletic functional group, ribosomal gene surveys are unlikely to detect specific responses of lignocellulolytic microorganisms.

Conversely, conventional methods commonly used in soil ecology to assess lignocellulolytic activities expressed in soil (e.g. measurement of extracellular enzymatic activities in soil extracts) provide only quantitative values that cannot be related to the diversity of saprotrophic species producing the corresponding enzymes.

An alternative approach is to target genes (functional markers) that encode enzymes involved in POM decomposition, and which are therefore directly linked to

the ecosystem process under investigation. Several PCR primers targeting fungal gene families encoding different lignocellulolytic enzymes have been designed (Maijala *et al.* 2003, Luis *et al.* 2004, Edwards *et al.* 2008, Bödeker *et al.* 2009, Kellner & Vandenbol 2010, Barbi *et al.* 2014, Kellner *et al.* 2014), and some of them successfully used to amplify the corresponding sequences from soil-extracted DNA and/or RNA (Luis *et al.* 2005, Edwards *et al.* 2008; Kellner & Vandenbol 2010, Weber *et al.* 2011, 2012, Baldrian *et al.* 2012, Barbi *et al.* 2014, Kellner *et al.* 2014; Mueller *et al.* 2014). This is the case for the fungal glycosyl hydrolase family 7 (GH7) encompassing enzymes active on cellulose, either cellobiohydrolases or endoglucanases (Edwards *et al.* 2008), fungal glycosyl hydrolase family 11 (GH11) with hemicellulolytic endoxylanase activities (Barbi *et al.* 2014) and Basidiomycota class II peroxidases potentially active on lignin (Bödeker *et al.* 2009, Barbi *et al.* 2014).

Besides the conflict between the use of phylogenetically or functionally-informative marker genes to assess microbial diversity, a recurrent discussion in microbial ecology regards the use and benefits of either environmental DNA or RNA as material to describe microbial communities. It is now generally accepted that RNA-based surveys of functional markers provide a more accurate information about taxa carrying out a specific activity at a given time than DNA-based studies, that probe all members of the community, whether or not they actively participate at the studied process at the time of sampling. For example, Baldrian *et al.* (2012) in a comparison of DNA- and RNA-based compositional surveys of GH7 (cellulases) sequences in forest litter and soil demonstrated that several of the most abundant fungal taxa in a DNA-based survey did not contribute significantly to the RNA pool (RNA-based survey).

To our knowledge, studies that have addressed the impact of environmental factors on fungal activities by using RNA-based approaches are very limited. Weber *et al.* (2012) assessed the combined impact of elevated atmospheric CO_2 and nitrogen fertilization in a loblolly pine (*Pinus taeda*) plantation on fungal GH7 gene diversity, used as a proxy to evaluate the diversity of the corresponding metabolically active cellulolytic fungal community. They observed no significant changes in richness and composition of the GH7 gene pool in response to the studied factors. However, the apparent lack of response by the GH7-producing fungi does not preclude that other components of the lignocellulolytic fungal community may be responding. Indeed (i) lignocellulolytic-encoding genes are not necessarily co-regulated and (ii) different saprotrophic or mycorrhizal fungal species each possess a specific repertoire of genes encoding enzymes active on POM (Floudas *et al.* 2012).This observation highlights the need to simultaneously characterize different fungal lignocellulolytic enzyme-encoding genes to better appreciate the impact of environmental changes on soil fungal community actively performing POM degradation.

To this aim, in order to elucidate the functional response of POM-degrading soil fungal community to an altered precipitation regime we simultaneously analyzed the diversity of three functional markers genes of the plant litter decomposition machinery by means of targeted metatranscriptomics coupled with high throughput sequencing. This allowed us to directly investigate *in-situ* the structure and changes of active lignocellulolytic fungal communities in soil samples collected in a Mediterranean holm oak (*Quercus ilex*) forest where a rainfall reduction experiment has been implemented since 2003. We also targeted the constitutively expressed eukaryotic gene Elongation Factor 1-alpha (EF1-α) to address if and how the overall fungal community (not limited to saprotrophic lignocellulolytic species, but including also symbiotic and pathogenic ones) responded to the simulated climate change.

Reduction in precipitation, by affecting soil water content, is susceptible to affect forest carbon allocation and turn-over. Indeed, in the Puéchabon forest throughfall reduction resulted in a significant and durable reduction in the leaf area index and therefore in the amount of foliar litter produced (Limousin *et al.* 2009). However, no significant effect on the increase in trunk diameter, taken as a proxy for tree above-ground biomass increase, could be observed. One possible explanation could be that a decrease in net photosynthesis resulted in changes in overall biomass allocation as illustrated by a significant decrease in flower and acorn production (investment in sexual reproduction) by trees subjected to rain reduction (Richard Joffre, personal communication). As above-ground vegetative biomass seemed not to be affected it can be hypothesized that a reduction occurs belowground. Such a belowground reduction may affect root biomass and/or C allocation to soil microorganisms.

3.3 Materials and Methods

3.3.1 *Study site and experimental design*

The study site is located 35 km north-west of Montpellier (southern France) in the Puéchabon State Forest (43°44’29’’N. 003°35’45’’E., 270 m a.s.l.; http://puechabon.cefe.cnrs.fr/site.htm). This Holm Oak (*Quercus ilex*) evergreen forest, established on Jurassic limestone, was managed for several centuries as a coppice and was last clear-cut in 1942. Understorey species compose a sparse shrubby layer dominated by *Buxus sempervirens, Phyllirea latifolia, Pistacia terebinthus* and *Juniperus oxycedrus*. Soil, classified as silty clay loam in texture (USDA classification) with a mull-type humus, is homogeneous in the 0-50 cm layer (39% clay, 26% sand, 35% silt). Climate is of the Mediterranean type with cold and wet winters and warm and dry summers, with rainfall occurring during autumn and winter (about 80% between September and April; Allard *et al.* 2008). The experimental site is equipped with a meteorological station, various probes and an Eddy-covariance flux tower allowing continuous measurements and estimations of

various climatic and functional parameters including soil water content or CO_2 and water exchanges at the atmosphere/vegetation interface.

In 2003, an experiment of reduction in throughfall precipitation (the precipitation falling through the forest canopy) was set up to evaluate the impact on forest ecosystem functioning of reductions in rainfall input as predicted for the Mediterranean area by climate change models for the current century (Mouillot *et al.* 2002, IPCC 2007). Three 20 m x 20 m plots, 150 m away one from each other (Limousin *et al.* 2009), were selected and within each of them two, 10 m x 10 m, areas were delimited. One was subjected to the "rainfall reduction" treatment and one was used as a control area receiving the natural rain input. In the "rainfall reduction" areas the net precipitation input to the soil was reduced by 29% (Limousin *et al.* 2008) compared to the control areas, by hanging, at a height between 0.8 m (lowest point) and 1.50 m (highest point) above ground, 14 m long x 0.19 m wide PVC gutters covering 33% of the ground surface area under the canopy. In the control areas, identical gutters were placed upside down so that the albedo and the micro-climate of the forest understorey were similarly affected in the treatment and control plots (Limousin *et al.* 2008).

Four sampling campaigns were conducted between 2010 and 2012. At each sampling campaign, in each of the plots (treatments and controls) ten soil cores of *ca* 200 cm^3 each (0-5 cm depth) were collected along a regular sampling grid, sieved (2 mm sieves) and mixed in equal amounts to constitute composite samples which were quickly frozen (sample list in Table S1). A total of 22 composite soil samples were thus collected and stored at -75°C prior to RNA extraction.

3.3.2 *Soil RNA extraction and cDNA synthesis*

For each sample, total RNA was extracted from at least 10g of soil following the protocol from Luis *et al.* 2005. Total soil RNA samples were treated with RNase free DNase I (Fermentas) to remove residual DNA contaminations and quantified by spectrophotometry (ND-1000 NanoDrop®, Thermo Scientific).

All eukaryotic cDNAs were synthesized and amplified using the Mint-2 cDNA synthesis kit (Evrogen) which allows the synthesis of full-length-enriched double stranded cDNA from poly-A mRNA present in total RNA samples. Synthesis was initiated using 2 µg of total soil RNA, and by using the 5'-end adapter PlugOligo-3M, and the 3'-end adapter CDS-4M as described in the Mint-2 kit instruction manual. Double stranded cDNAs (ds cDNAs) were amplified by PCR using the PCR primer M1 (AAGCAGTGGTATCAACGCAGAGT) which binds at both ends of each of the neo-synthesized cDNA. The number of optimal PCR cycles for the synthesis of the ds cDNA was evaluated for each sample used and ranged between 18 and 22.

3.3.3 *Gene-specific PCR amplifications and sequencing*

Pre-amplified ds cDNAs were used as templates for PCR reactions targeting three major fungal gene families coding enzymes involved in the degradation of plant cell-wall polymers. The Glycoside Hydrolase 11 (GH11, according to the CAZy database, http://www.cazy.org/) encodes *endo*xylanases (EC. 3.2.1.8) active on hemicelluloses, the GH7 family encodes cellulases (either cellobiohydrolases, EC. 3.2.1.176 or *endo*-β-1,4-glucanases, EC 3.2.1.73) and the AA2 family encodes Basidiomycota-class II peroxidases (EC1.11.1.-; including so-called generic, manganese, lignin and versatile peroxidases) active on lignin and other aromatic compounds. In parallel, the elongation factor 1-α (EF1-α) gene, a housekeeping and constitutively-expressed eukaryotic gene, was also amplified to assess the diversity of the overall active soil eukaryotic biota. Primer sequences (described in Rehner and Buckley 2005, Edwards *et al.* 2008 and Barbi *et al.* 2014) and sizes of the expected amplified fragments are given in Table 1. PCR reaction mixes contained 1 µl of template cDNA, 2.5 µl of 10X PCR buffer without Mg (Invitrogen), 1.5 mM $MgCl_2$, 0.8 mM of each dNTP, 0.5 µM of each primer and 1.25 U of DNA polymerase (a 24:1 mix of Invitrogen *Taq* DNA polymerase and Biorad iProof polymerase). After an initial denaturation at 94°C for 3 min, amplification proceeded for 45 (GH11), 35 (GH7 and AA2) or 25 cycles (EF1-α) comprising 45 s of denaturation at 94°C, 45 s of primer annealing at 48°C (GH7) or 50°C (GH11, AA2 and EF1-α primers) and 2 min of synthesis at 72°C. Amplifications were terminated by a 10 min final extension at 72°C.

Five separate PCR tubes were prepared for each gene and run in parallel for each ds cDNA sample. Each PCR product was controlled by running 5 µl on 1.5% agarose gels before pooling the five replicates. GH11, GH7 and EF1-α PCR products were directly purified using the QIAquick PCR purification kit (Qiagen). The AA2 PCR products were first run on an agarose gel and the PCR fragment of the expected size (Table 1) was extracted and purified using the QIAquick Gel extraction kit (Qiagen). After quantification by fluorimetry (Qubit 2.0 fluorometer, Invitrogen), a mix of the four different PCR products was prepared for each of the 22 samples and a paired-end sequencing (2x250bp) was carried out on an Illumina MiSeq sequencer by FASTERIS (Switzerland).

3.3.4 *Data analysis*

Each Illumina paired-end read was assembled with PEAR v.0.9.2 (Zhang *et al.* 2014) and assembled sequences containing unidentified nucleotide positions ("N") were discarded. After primer detection, sequences from each of the four gene families were demultiplexed, and subsequent analyses were carried out independently for each gene family. Sequences were processed with MOTHUR v. 1.33.0 (Schloss *et al.* 2009), chimeras were detected with UCHIME (Edgar *et al.*, 2011) and removed. As GH7 PCR products were approx. 500 bp in length, we limited our analysis of these

sequences to the first 210 nucleotides of the reads bordered by the GH7 forward primer (Barbi *et al.* 2014).

Nucleotide sequences were clustered into "Operational Functional Units" (OFUs) using a 93% identity threshold for AA2, a 95% id. threshold for GH11 and GH7 (Barbi *et al.* 2014), and a 96% id. threshold for EF1-α. The most abundant sequence from each cluster was selected as its representative sequence and translated in the corresponding amino acidic sequences with FRAMEBOT (Wang *et al.* 2013). Clusters with sequences containing an in-frame stop codon were discarded.

Since EF1-α sequences were generated using PCR primers that were not specific for fungal sequences, non-fungal sequences were identified and discarded by using a BLASTx search against the GenBank non-redundant protein database. Sequences for which the first five best BLAST hits were fungal sequences were considered as being of fungal origin.

Because different sampling intensities may lead to bias and erroneous conclusions (Gihring *et al.* 2012), the data were analysed after equalization of the sampling effort. For each gene, all 22 samples were randomly subsampled to the same sequencing depth (*i.e.* rarefied to the sample with the lowest number of sequences, Table 2). Richness (observed number of clusters and Chao1 index) and Shannon's diversity indices were then calculated and compared using the Kruskal-Wallis non parametric ANOVA test.

Because the low frequency clusters may have disproportionate effects on ordination and multivariate analyses, these statistical analyses were performed on datasets from which rare clusters had been omitted. Rare clusters (encompassing the so-called singletons) were defined as clusters present in less than two of either the control or treatment (rain exclusion) plots for each sampling date.

Non-metric multidimensional scaling (NMDS) ordination was performed using both Jaccard and Bray-Curtis distance matrices, based on OFU presence/absence and abundance (number of sequences included in the cluster following subsampling), respectively. PERMANOVA analysis was also carried out using both Jaccard and Bray-Curtis distance matrices to investigate the relationship between community composition and the different factors considered in the study (treatment, plot, sampling date and season).

PERMANOVA was also carried out on Generalized UniFrac (Chen *et al.* 2012) distance matrices, calculated for each gene to explore differences in phylogenetic diversity among samples. Generalized UniFrac distances were calculated on phylogenetic trees constructed with amino acid sequences corresponding to the translation of the representative sequence of each cluster. Amino acid sequences were aligned with MUSCLE (Edgar, 2004) and phylogenetic trees were generated

with PhyML 3.0 using the WAG substitution model as implemented in SeaView v. 4 (Gouy *et al.*, 2010).

Variance partitioning was also carried out to estimate the amount of variance in single gene communities that could be explained by each of the studied factors.

Statistical analyses were carried out using the *vegan* (Oksanen *et al.* 2013), *phyloseq* (McMurdie & Holmes 2013) and *GUniFrac* packages in R environment (R Core Team, 2014)

Indicator sequence clusters for different factors were identified using the R package *indicspecies* (De Cáceres & Legendre, 2009). Only indicator clusters having both sensitivity (the probability that a surveyed sample in which the indicator cluster is present belongs to the target sample group) and specificity (the probability of finding the indicator cluster in a given sample belonging to the target sample group) values higher than 0.75 were considered. Amino acid sequences corresponding to the selected indicator clusters were aligned to reference sequences retrieved from public databases (Joint Genome Institute http://jgi.doe.gov/, GenBank (http://www.ncbi.nlm.nih.gov/genbank/). These reference sequences (maximum five sequences for each of the query sequences) were selected as those with the best BLASTx hits in the mentioned databases. Phylogenetic trees were built as described above. Phylogenetic trees were drawn and adjusted using MEGA6 (Tamura *et al.*, 2013).

3.4 Results

3.4.1 *Rain exclusion effects and soil sampling*

Rain exclusion resulted in a consistent and continuous deficit in soil water content compared to control conditions across all years (Figure 1). Rain exclusion, however, had no evident effect on other measured soil parameters such as soil temperature (Figure S1) or overall soil composition (Table S2).

Soil samples were taken in spring and autumn, two seasons with significant amounts of rainfall and mild temperatures but which differ from each other by gross ecosystem productivity. In spring, the balance between Gross Primary Production (GPP) and ecosystem respiration (R_{ECO}) results in a positive carbon accumulation at the ecosystem scale (translated as negative Net Ecosystems Exchange (NEE) values; Figure 2). In autumn, this balance is in favour of respiration and results in carbon loss at the ecosystem scale (positive NEE values, Figure 2). Samplings were performed after "long" periods of rain and with soil temperatures above 5-13 °C, which should favour soil microbial activities in both control and rain-exclusion plots (Figure S2).

Gene family/Enzymes[1]	Primer sets	Primer sequences (5'-3')	PCR fragment lengths (bp)	Targeted fungal groups	References
GH7/ Cellulases	fungcbhI-F fungcbhI-R	ACCAAYTGCTAYACIRGYAA GCYTCCCAIATRTCCATC	515	Basidiomycota Ascomycota	Edwards *et al.* 2008
GH11/ Endoxylanases	fungGH11-F fungGH11-R	GGVAAGGGITGGAAYCCNGG TGKCGRACIGACCARTAYTG	281	Basidiomycota Ascomycota	Barbi *et al.* 2014
AA2/ Class II peroxidases	basidioPerox-F basidioPerox-R	GGYGGIGGIGCBGAYGGYTC GGRGTIGAGTCRAANGG	398	Basidiomycota	Barbi *et al.* 2014
Elongation factor 1-α	EF1-1577F Efgr	CARGAYGTBTACAAGATYGGTGG GCAATGTGGGCRGTRTGRCARTC	350	All Eukarya	Rehner & Buckley 2005

Table 1. Primer sets used to amplify the targeted genes, [1]according to the CAZY database (http://www.cazy.org/, Lombard *et al.* 2014).

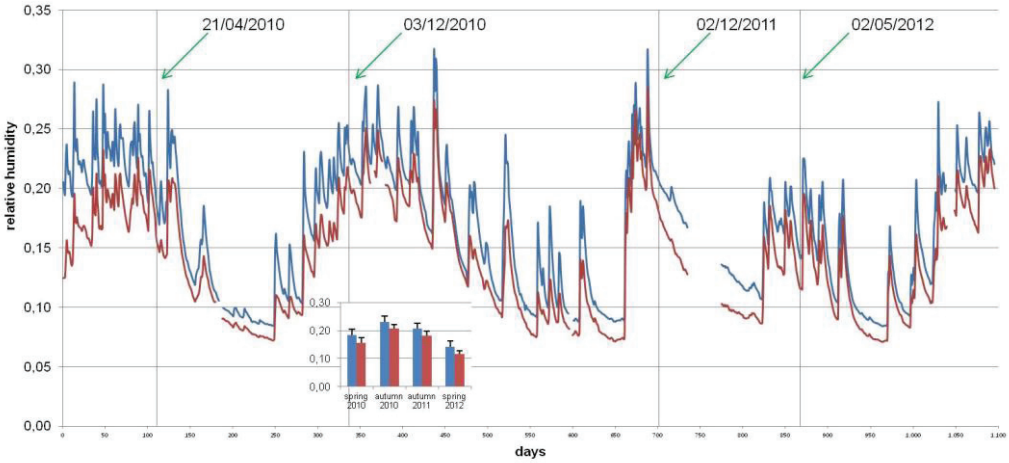

Figure 1. Variations in mean soil relative humidity (cm^3 of water per cm^3 of soil) of the control (blue line) and rainfall reduction (red line) plots (from January 2010 to December 2012). The four indicated dates correspond to the soil sampling dates. Insert: values for the four sampling dates.

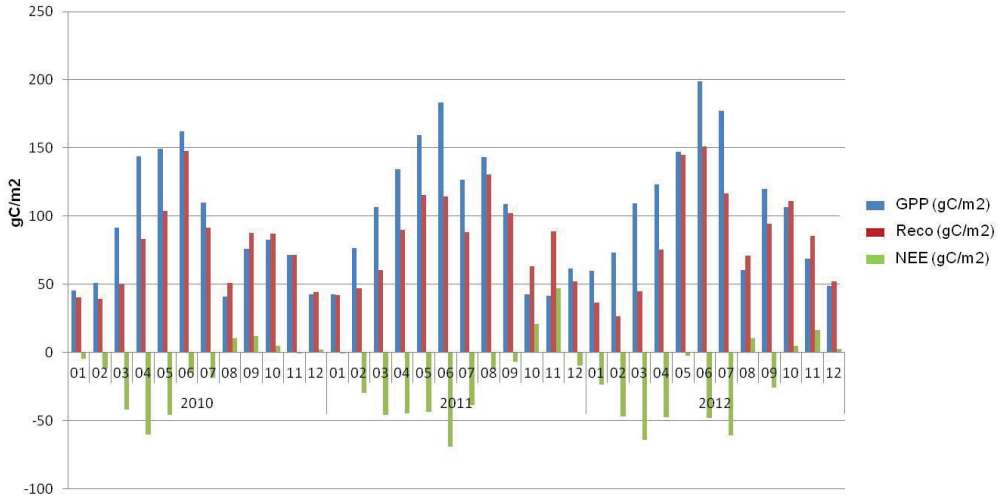

Figure 2. Monthly Ecosystem Carbon balance GPP (Gross Primary Production), Reco (Ecosystem Respiration), NEE (Net Ecosystem Exchange). Negative NEE values indicate net CO_2 uptake by the ecosystems, and positive values indicate net release of CO_2 to the atmosphere. Measurements were performed with the Eddy-covariance flux tower.

3.4.2 *Sequences and clusters*

RNA was successfully extracted from the 22 soil composite samples and ds cDNA synthesized from eukaryotic mRNAs. GH11, GH7, AA2 and EF1-α gene fragments were amplified from all the samples and subjected to high-throughput Illumina MiSeq sequencing.

A total of 1,769,740 sequences were recovered after primer filtering and chimera removal and were used for the clustering at different levels of identity, depending on the targeted gene. The percentage of sequences belonging to each gene family was 55%, 4%, 11%, and 30% for the GH11, GH7, AA2 and EF1-α families respectively (Table 2). The representative sequences of each sequence cluster (hereafter referred to as "Operational Functional Units", OFUs) were further translated into the corresponding amino acid sequences and the OFUs characterized by a sequence containing a premature in-frame stop codon were discarded. Depending on the gene family, this filter resulted in the elimination of between 4% and 8% of the total OFUs (representing between 0.02 and 0.8% of the total sequences) (Table 2). Regarding the EF1-α gene, out of the 11,655 OFUs obtained, about half of them (5016 OFUs) were retained as being of fungal origin following BLASTx analyses.

The OFU amino acid sequences were all homologous to fungal lignocellulolytic enzyme sequences of the corresponding CAZyme family already deposited in public databases, with variable percentages of identity to the most similar public sequence, depending on the functional gene (Fig. 3). The higher proportion of GH7 sequences with high percentage of identity to known sequences is certainly due to the higher number of sequences deposited in public databases for this gene family compared to the others and, moreover, to the higher number of sequences from uncultured soil eukaryotes for this gene.

Gene	Clustering identity threshold	No. of clusters / No. of seq.	No. of clusters / No. of seq. removed*	No. sequences per sample after subsampling	No. of clusters after subsampling	No. of clusters / sequences after frequency selection
GH11	95%	3361/972,980	280 (8.3%)/1148 (0.12%)	16,554	1544	188/306,853
GH7	95%	1334/65,789	87 (6.5%)/119 (0.2%)	344	432	40/4327
AA2	93%	534/191,176	26 (5%)/34 (0.02%)	826	233	45/14,097
EF1-α all	96%	12,116/539,795	461 (4%)/4364 (0.8%)	10,047	7835	---
Fungal EF1-α	96%	5016/262,652	---	4705	3061	627/89,926

Table 2. Numbers of clusters (each representing an OFU, Operational Functional Unit) and sequences for each gene analyzed. *representative sequences with a premature stop codon.

% of identity	*GH11*	*GH7*	*AA2*	*EF1-α*
Min.	51.06	50	50	66.99
1st Qu.	77.50	75	63.41	91.26
Median	82.28	82.61	68.49	95.15
Mean	8.50	81.82	69.91	93.05
3rdQu.	86.08	89.86	76.09	97.09
Max.	100	100	100	100

Table 3. Distribution of the percentage of identity between the amino acid sequences obtained in the present work (before subsampling) and their most similar sequences present in public databases.

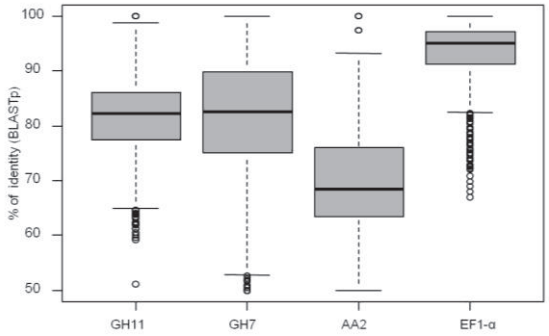

Figure 3. Box-plots of the percentages of identity between the amino acid sequences obtained in the present study and the most similar sequences present in public databases (as in June 2013). Graphs are presented for each four studied gene families.

The mean percentage of amino acid sequence similarity to known sequences for the EF1-α is higher (93%; Fig.3, Table 3) than those calculated for the other genes. For several sequences it even reached 100% identity. However, the amplified EF1-α gene fragment is not the most suitable one to assign sequences at the species level, since distinct fungal species can share identical amino acid sequences.

Rarefaction curves (Fig. S3) indicated that none of the genes reached the sampling saturation in all 22 samples; we thus rarefied the samples to the same sequencing depth to reduce the effects of unequal sampling (16,554, 344, 826 and 4705 sequences *per* sample were obtained for the GH11, GH7, AA2 and fungal genes respectively). These sequences fall into 188 (GH11), 40 (GH7), 45 (AA2) and 627 (EF1-α) clusters after the removal of rare OFUs.

Gene families consistently differed from one another with respect to OFU richness (both observed no. of OFUs and estimated Chao1), with EF1-α > GH11 > GH7 > AA2 in almost all samples (Table 4).

Richness and Shannon diversity indices, calculated for each gene after sub-sampling, did not significantly differ between treatments (rainfall reduction *vs* control), between treatments at each sampling date, between different sampling dates and between plots (Kruskal-Wallis/Mann-Whitney pairwise comparisons; Table 4). Only the GH11 OFU richness of soil samples collected in autumn 2010 was significantly higher than the GH11 OFU richness of soil samples collected in spring 2012 (P=0.0098, Mann-Whitney pairwise comparisons).

Sample groups	GH11						GH7					
	R mean	***R SD***	**Chao1 mean**	**Chao1 SD**	**H mean**	***H SD***	**R mean**	**R SD**	**Chao1 mean**	***Chao1 SD***	**H mean**	***H SD***
rainfull reduction	153.3	*63.2*	327.86	*166.7791*	2.6	*0.8*	38.7	*12.1*	71.56	*26.73*	2.7	*0.5*
control	122.5	*53.3*	261.2104	*125.2229*	2.6	*0.7*	32.6	*11.0*	53.00	*25.16*	2.3	*0.6*
Dry_spring 2010	195.5	*43.1*	438.17	*93.46*	3.2	*0.4*	42.5	*4.9*	112.85	*16.48*	2.9	*0.2*
Control_spring_2010	133.0	*9.9*	380.43	*62.68*	2.6	*0.7*	28.0	*7.1*	35.67	*4.71*	2.3	*0.2*
Dry_autumn 2010	191.7	*21.7*	497.99	*84.19*	3.1	*0.7*	34.3	*4.0*	57.13	*17.94*	2.4	*0.3*
Control_autumn 2010	165.7	*42.3*	345.48	*113.90*	3.2	*0.5*	38.3	*10.7*	75.15	*27.51*	2.6	*0.5*
Dry_autumn 2011	114.3	*44.4*	216.17	*64.98*	2.5	*0.5*	40.0	*8.7*	69.83	*10.98*	2.8	*0.2*
Control_autumn 2011	105.7	*67.6*	188.25	*81.87*	2.5	*0.9*	32.7	*7.5*	50.12	*13.02*	2.3	*0.7*
Dry_spring_2012	125.7	*95.8*	195.88	*151.56*	1.7	*0.8*	39.3	*24.1*	60.19	*27.92*	2.6	*0.9*
Control_spring_2012	89.0	*52.9*	170.41	*107.46*	2.1	*0.5*	30.0	*18.2*	45.30	*33.01*	2.2	*1.1*
Spring 2010	164.3	*44.2*	**a**409.30	*73.02*	**a** 2.9	*0.6*	35.3	*9.7*	74.26	*45.65*	2.6	*0.4*
Autumn 2010	178.7	*33.3*	**a**421.74	*122.48*	**a.b** 3.2	*0.6*	36.3	*7.6*	66.14	*23.00*	2.5	*0.4*
Autumn 2011	110.0	*51.4*	**b**202.21	*67.85*	**a** 2.5	*0.6*	36.3	*8.3*	59.98	*15.25*	2.6	*0.5*
Spring 2012	107.3	*72.1*	**c**183.15	*118.33*	**a.c** 1.9	*0.6*	34.7	*19.8*	52.75	*28.53*	2.4	*0.9*
Plot1	149.9	*48.6*	325.74	*129.86*	2.8	*0.6*	37.1	*10.3*	60.91	*26.89*	2.6	*0.4*
Plot2	143.4	*50.0*	296.95	*134.59*	2.6	*0.9*	34.6	*6.0*	65.05	*33.98*	2.5	*0.3*
Plot3	114.5	*83.5*	249.72	*197.31*	2.2	*0.6*	35.2	*19.3*	74.50	*15.42*	2.3	*1.0*

Table 4. Alpha diversity descriptors for different groups of samples. R: richness (total no. of observed clusters); Chao1 index; H, Shannon diversity index; SD, standard deviation. Significant differences within a given sample group were tested with non-parametric Kruskal-Wallis/Mann-Whitney pairwise comparisons. Different bold letters indicates significant differences ($P<0.05$).

Sample groups	AA2						Fungal EF1-α					
	R mean	*R SD*	Chao1 mean	*Chao1 SD*	H mean	*H SD*	R mean	*R SD*	Chao1 mean	*Chao1 SD*	H mean	*H SD*
rainfull reduction	25.4	*8.6*	32.23	*12.95*	1.9	*0.3*	464.9	*104.8*	677.76	*184.75*	4.8	*0.4*
control	26.2	*8.9*	42.42	*28.89*	1.9	*0.5*	384.3	*95.9*	574.95	*150.34*	4.7	*0.4*
Dry_spring 2010	27.0	*27.0*	33.08	*18.50*	2.1	*0.4*	528.5	*74.2*	778.34	*114.29*	4.9	*0.8*
Control_spring_2010	35.5	*35.5*	81.25	*58.34*	2.0	*0.0*	366.0	*4.2*	475.44	*17.87*	4.3	*0.5*
Dry_autumn 2010	19.0	*4.4*	21.49	*6.63*	1.6	*0.1*	451.0	*51.4*	619.10	*69.14*	5.0	*0.3*
Control_autumn 2010	21.0	*8.7*	29.38	*11.26*	1.6	*0.4*	386.7	*114.2*	575.88	*162.71*	4.9	*0.2*
Dry_autumn 2011	23.3	*4.2*	28.39	*6.82*	1.8	*0.3*	427.0	*73.5*	613.42	*139.40*	4.7	*0.5*
Control_autumn 2011	24.7	*4.6*	38.17	*11.54*	2.1	*0.1*	430.7	*29.9*	676.68	*91.35*	4.8	*0.2*
Dry_spring_2012	32.7	*9.0*	46.22	*9.37*	2.0	*0.4*	474.3	*193.6*	733.71	*337.26*	4.8	*0.4*
Control_spring_2012	26.7	*12.7*	33.83	*17.61*	1.8	*0.8*	347.7	*162.5*	538.64	*225.02*	4.6	*0.6*
Spring 2010	31.3	*9.6*	57.17	*44.96*	2.1	*0.2*	447.3	*105.5*	626.89	*187.20*	4.6	*0.5*
Autumn 2010	20.0	*6.3*	25.44	*9.33*	1.6	*0.3*	418.8	*86.7*	597.49	*114.29*	4.9	*0.2*
Autumn 2011	24.0	*4.0*	33.28	*10.02*	2.0	*0.2*	428.8	*50.2*	645.05	*110.96*	4.7	*0.3*
Spring 2012	29.7	*10.4*	40.03	*14.32*	1.9	*0.6*	411.0	*174.3*	636.18	*277.79*	4.7	*0.5*
Plot1	25.6	*7.6*	41.21	*34.00*	3.1	*1.1*	432.1	*90.2*	627.72	*149.02*	4.8	*0.5*
Plot2	28.8	*6.5*	39.03	*9.62*	3.2	*0.6*	432.6	*48.7*	631.90	*106.76*	4.7	*0.3*
Plot3	22.0	*11.6*	29.87	*16.58*	2.7	*0.7*	403.8	*179.1*	617.15	*279.11*	4.8	*0.4*

Table 4. Alpha diversity descriptors for different groups of samples. R: richness (total no. of observed clusters); Chao1 index; H, Shannon diversity index; SD, standard deviation. Significant differences within a given sample group were tested with non-parametric Kruskal-Wallis/Mann-Whitney pairwise comparisons. Different bold letters indicates significant differences ($P<0.05$).

GH11

	Bray-Curtis distance		Jaccard distance		GUniFrac d=0.5	
Factor	***r²***	***P-value***	***r²***	***P-value***	***r²***	***P-value***
Treatment	0.06275	0.1419	**0.06735**	**0.02797**	0.054656	0.263
sampling_date	**0.23762**	**0.000999**	**0.22773**	**0.000999**	**0.23978**	**0.002**
season	**0.11286**	**0.000999**	**0.0797**	**0.003996**	**0.13528**	**0.002**
plot	0.10989	0.2058	0.09146	0.5894	0.09058	0.548
Treatment*sampling_date	0.11714	0.571429	**0.15418**	**0.014985**	0.12157	0.532
Treatment*season	0.04687	0.372627	0.04918	0.274725	0.02994	0.848
Treatment*plot	0.09655	0.3746	0.09699	0.35764	0.08719	0.575
plot*sampling_date	0.20778	0.420579	0.18386	0.984016	0.21797	0.309
plot*season	0.08551	0.494505	0.08642	0.677323	0.08816	0.396

AA2

	Bray-Curtis distance		Jaccard distance		GUniFrac d=0.5	
Factor	***r²***	***P-value***	***r²***	***P-value***	***r²***	***P-value***
Treatment	0.05011	0.3606	0.04973	0.3856	0.7239	0.127
sampling_date	0.27675	**0.003996**	**0.27218**	**0.000999**	**0.30166**	**0.002**
season	0.14545	**0.000999**	**0.14996**	**0.000999**	**0.15724**	**0.001**
plot	0.11945	0.1768	0.07353	0.8372	0.09131	0.496
Treatment*sampling_date	0.12068	0.42957	**0.1815**	**0.012987**	0.12829	0.239
Treatment*season	0.04153	0.447552	0.0547	0.166833	0.04152	0.394
Treatment*plot	0.07613	0.7363	0.06212	0.957	0.06328	0.872
plot*sampling_date	0.1715	0.693307	0.17013	0.963507	0.1626	0.698
plot*season	0.07105	0.645355	0.06067	0.09091	0.69703	0.813

GH7

	Bray-Curtis distance		Jaccard distance		GUniFrac d=0.5	
Factor	***r²***	***P-value***	***r²***	***P-value***	***r²***	***P-value***
Treatment	0.04246	0.5974	0.04117	0.6503	0.03182	0.837
Sampling date	**0.22066**	**0.001998**	**0.20974**	**0.003996**	**0.22862**	**0.006**
season	0.06815	0.1029	0.07027	0.08891	**0.09737**	**0.007**
plot	0.11741	0.1419	0.10265	0.3387	0.10599	0.291
Treatment*sampling date	0.13857	0.352647	**0.18464**	**0.011988**	0.15033	0.193
Treatment*season	0.03425	0.8062	0.04699	0.45055	0.03814	0.639
Treatment*plot	0.06904	0.8941	0.1014	0.3157	0.0825	0.677
plot*sampling date	**0.24578**	**0.047952**	0.21326	0.515485	0.23907	0.112
plot*season	0.09402	0.3966	0.08079	0.67932	0.08696	0.516

Fungal EF1-α

	Bray-Curtis distance		**Jaccard distance**		**GUniFrac d=0.5**	
Factor	***r²***	***P-value***	***r²***	***P-value***	***r²***	***P-value***
Treatment	0.05567	0.1738	**0.07283**	**0.000999**	0.05745	0.151
Sampling date	**0.18531**	**0.004995**	**0.18091**	**0.004995**	**0.22791**	**0.001**
season	**0.07059**	**0.01698**	**0.06281**	**0.02198**	0.06274	0.063
plot	**0.15669**	**0.000999**	**0.13419**	**0.000999**	**0.12252**	**0.028**
Treatment*sampling date	0.10555	0.995	0.13085	0.521479	0.10382	0.977
Treatment*season	0.03452	0.92507	0.04534	0.458541	0.0324	0.959
Treatment*plot	**0.11229**	**0.014985**	**0.11143**	**0.006993**	0.09896	0.263
plot*sampling date	0.17406	0.998002	0.18615	0.999001	0.19397	0.763
plot*season	0.06539	0.985015	0.07524	0.962038	0.08241	0.736

Table 5. PERMANOVA (1000 permutations) for the GH11, GH7, AA2 and fungal EF1-α datasets based on Bray-Curtis, Jaccard and GUniFrac distances. Significant results (P<0.05) are in bold type.

	Fungal EF1-α Plot1			
	Bray-Curtis distance		**Jaccard distance**	
Factor	***r²***	***P-value***	***r²***	***P-value***
Treatment	0.1482	0.4545	**0.17263**	**0.04196**
Sampling date	0.42564	0.4995	0.42117	0.5594
season	0.13295	0.7273	0.13908	0.5455
Treatment*sampling date	0.1482	1	0.40620	1
Treatment*season	0.14013	0.6194	0.14691	0.29071
	Fungal EF1-α Plot2			
	Bray-Curtis distance		**Jaccard distance**	
Factor	***r²***	***P-value***	***r²***	***P-value***
Treatment	0.18284	0.06194	**0.17424**	**0.03097**
Sampling date	0.43731	0.4036	0.45529	0.1229
season	0.15449	0.2907	0.15063	0.3646
Treatment*sampling date	0.37985	1	0.37047	1
Treatment*season	0.11168	0.86913	0.11019	0.97602

Table 6. PERMANOVA (1000 permutations) for the fungal EF1-α dataset performed separately for each of the three sampled plots based on Bray-Curtis and Jaccard distances. Significant results (P<0.05) are in bold type.

	Fungal EF1-α Plot3			
	Bray-Curtis distance		Jaccard distance	
Factor	r^2	***P-value***	r^2	***P-value***
Treatment	0.24509	0.09291	0.26113	0.0969
Sampling date	0.39256	0.4446	0.37658	0.6663
season	0.20179	0.4096	0.1888	0.5015
Treatment*sampling date	0.36235	1	0.37658	1
Treatment*season	0.18104	0.58142	0.19573	0.41658

Table 6. PERMANOVA (1000 permutations) for the fungal EF1-α dataset performed separately for each of the three sampled plots based on Bray-Curtis and Jaccard distances. Significant results (P<0.05) are in bold type.

3.4.3 *Factors impacting functional gene community composition*

PERMANOVA analyses based on Bray-Curtis distances indicated that both the sampling date and the sampling season had a significant impact (P<0.004) on community composition for all but one of the four analyzed genes (only season had no significant effect on the GH7 "communities"), while the treatment (rainfall reduction *versus* control) did not (Table 5). A similar result pointing to a significant effect of sampling date and season in shaping community compositions was obtained in PERMANOVAs comparing phylogenetic GUniFrac distances (Table 5).

A significant effect of the treatment (rainfall reduction) was, however detected in the case of the GH11 gene family, when PERMANOVA was carried out using Jaccard distances (which take into account OFU presence/absence) (Table 5). However, non-metric multidimensional scaling ordination analysis (NMDS) (Fig. 4) computed using GH11 Jaccard dissimilarity distances indicated that treatment had a less pronounced effect than sampling date (higher segregation of samples collected in different sampling dates).

As sampling date and season are not independent variables, variance partitioning analysis was carried out twice, using either sampling date or the season as the temporal variables along with the other factors (Fig. 5 and Fig. S5). This analysis showed that the highest percentage of variance explained by the factors investigated (sampling dates, season, plots and treatment) was 25% (AA2 community). A greater proportion of variation in soil OFU communities was always explained by the sampling day than by the treatment (Fig. 5 A-C). The proportion of variation explained by the sampling day is also higher than the proportion explained by the sampling plot (spatial heterogeneity). A large amount of OFU community structure variance could not be explained, indicating that other factors which were not considered in the present study are important drivers of soil OFU communities. The

same result was obtained when considering the season (instead of sampling date) as the temporal variable (Fig. S5 A-C).

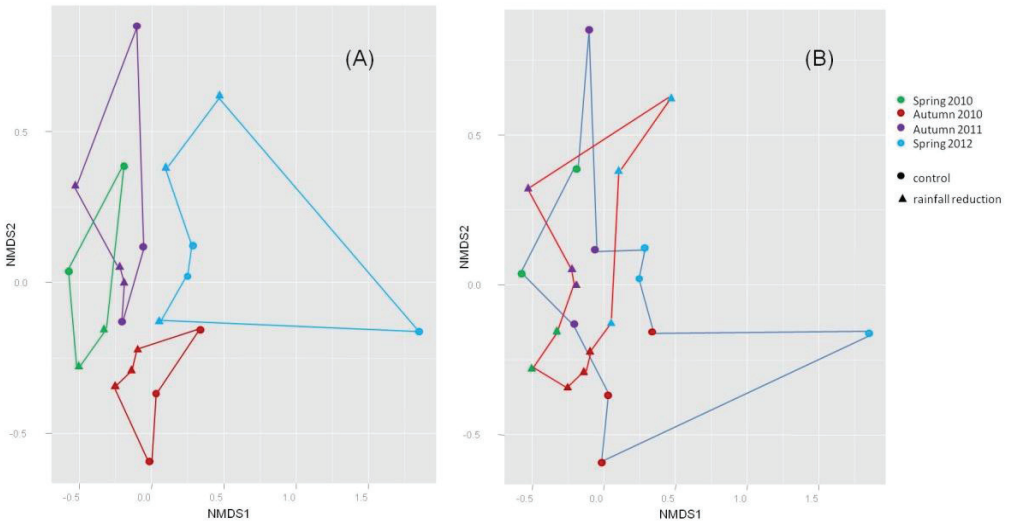

Figure 4. Non-metric multidimensional scaling (NMDS) plots of the Jaccard matrix of the GH11 OFUs (sequence clusters). (A) Samples collected on the same day and (B) samples subjected to the same treatment (control/rainfall exclusion) are connected by lines of the same colour.

3.4.4 *Overall active fungal community (fungal EF1-α community)*

A significant spatial effect (in addition to significant effects of both season and sampling date) was observed for EF1-α gene diversity (taken as a proxy of the diversity of the total active fungal community), highlighting differences in fungal community composition between the three sampled plots. These effects were observed with PERMANOVA analyses using either Bray-Curtis, Jaccard or GUniFrac distances (only seasonal effects were not significant in the analysis using GUniFrac distances; Table 5).

A significant treatment effect using the Jaccard distances, as well as significant treatment*plot interaction effects using both Jaccard distances and Bray-Curtis indices, were observed (Table 5). We therefore repeated the PERMANOVA analyses independently for each of the three plots and identified significant treatment effects for two of them (plots 1 and 2), but not for the third one, using Jaccard distance matrices (Table 6).

As for GH11, GH7 and AA2 OFU community composition, a large amount of fungal EF1-α community structure variance could not be explained by the factors we have considered. Altogether the factors considered explained 14.6% of the total variance. Contrasting with the other genes, the greater proportion of variation in

overall active fungal communities was explained by the sampling plot rather than by the treatment and sampling date/season, which explained lower and comparable amounts of variance (Fig. 5 d, Fig. S5 d).

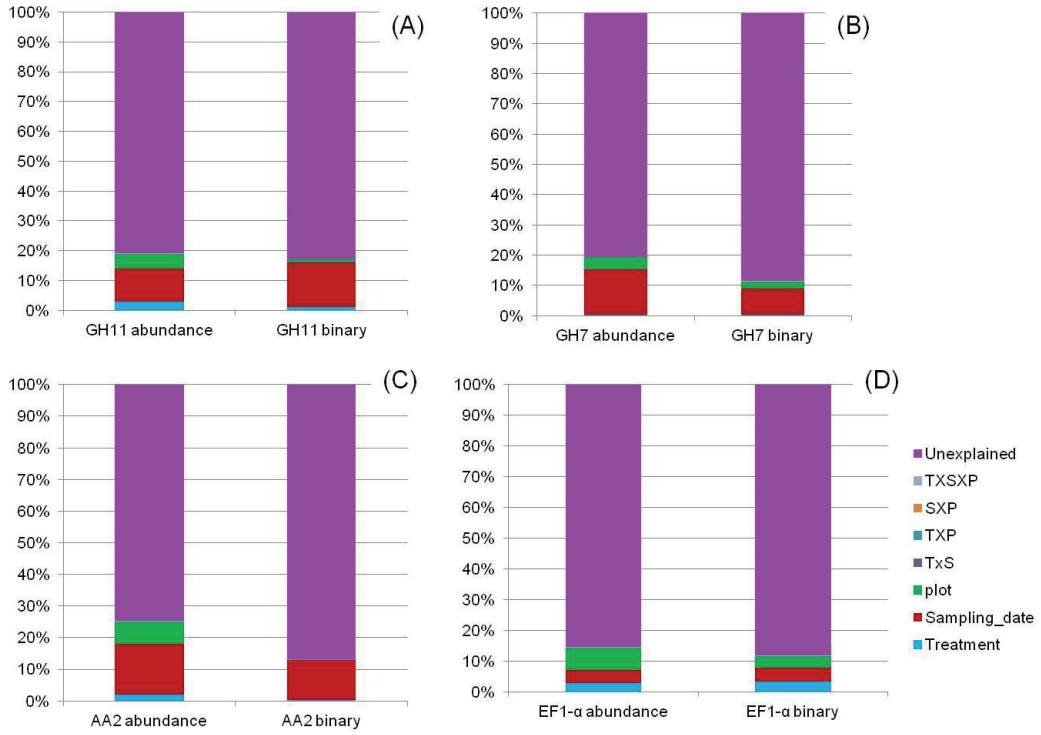

Figure 5. Variance partitioning, displayed as bar plots, depicting the relative contribution of treatment (T), sampling date (S) and plot (P) factors to total community variance. Variance partitioning was calculated using abundance data (after Hellinger transformation) and binary (presence/absence) data. (A) GH11, (B) GH7, (C) AA2, (D) Fungal EF1-α community.

3.4.5 *Indicator OFU analysis*

We performed an indicator OFU analysis in an attempt to identify OFUs associated to one or several of the environmental factors under study. By using stringent criteria (specificity and sensitivity values both above 0.75), analysing the whole dataset (22 samples) we could identify 29 EF1-α and 15 GH11 OFUs, but no GH7 or AA2 OFUs, as being indicators of a specific sampling date and season. Eight EF1-α and one GH11 OFUs were identified as indicators of either the rainfall reduction treatment or the control (Table 7). However, when the analysis was repeated separately on each of the four sampling date datasets, no OFU was found as being indicator of either the rainfall reduction treatment or the control.

Aminoacid sequences representative of each of the different indicator OFUs were used to build EF1-α and GH11 phylogenetic trees (Figs 6-7), along with two to five reference sequences selected as the most similar to the OFU sequences in BLASTp searches in the GenBank and Mycocosm databases. Indicator OFU sequences for a specific factor (either sampling date, season, control or rainfall reduction) did not cluster together and obviously belonged to different fungal phyla (e.g. Glomeromycota, Basidiomycota or Ascomycota). Absence of highly similar reference sequences in public databases precluded more precise taxonomic affiliation of the indicator OFU sequences.

OFU	*group*	*P-value*
EF_OFU233	control	0.013
EF_OFU84	control	0.006
GH11_OFU47	control	0.009
EF_OFU60	rainfall reduction	0.032
EF_OFU112	rainfall reduction	0.014
EF_OFU125	rainfall reduction	0.007
EF_OFU164	rainfall reduction	0.014
EF_OFU179	rainfall reduction	0.031
EF_OFU218	rainfall reduction	0.041
EF_OFU74	autumn	0.038
EF_OFU82	autumn	0.003
EF_OFU99	autumn	0.012
EF_OFU102	autumn	0.032
EF_OFU106	autumn	0.02
EF_OFU188	autumn	0.031
EF_OFU200	autumn	0.002
EF_OFU206	autumn	0.009
GH11_OFU24	autumn	0.03
GH11_OFU40	autumn	0.041
EF_OFU143	spring	0.002
EF_OFU148	spring	0.026
EF_OFU189	spring	0.003
EF_OFU211	spring	0.025
EF_OFU244	spring	0.002
EF_OFU251	spring	0.026
EF_OFU471	spring	0.022
EF_OFU502	spring	0.028
GH11_OFU5	spring	0.026
GH11_OFU75	spring	0.003
EF_OFU436	spring 2010	0.017
EF_OFU493	spring 2010	0.001
EF_OFU494	spring 2010	0.028
EF_OFU495	spring 2010	0.033
EF_OFU534	spring 2010	0.026
EF_OFU555	spring 2010	0.014
EF_OFU570	spring 2010	0.018
GH11_OFU60	spring 2010	0.002
GH11_OFU62	spring 2010	0.003
GH11_OFU75	spring 2010	0.003
GH11_OFU80	spring 2010	0.001
GH11_OFU84	spring 2010	0.003
GH11_OFU88	spring 2010	0.003
GH11_OFU97	spring 2010	0.023
GH11_OFU102	spring 2010	0.005
EF_OFU1	autumn 2010	0.003
EF_OFU194	autumn 2010	0.017
EF_OFU247	autumn 2010	0.001
GH11_OFU49	autumn 2010	0.027
EF_OFU21	autumn 2011	0.004
EF_OFU2	spring 2012	0.005
EF_OFU16	spring 2012	0.005
EF_OFU244	spring 2012	0.01
GH11_OFU23	spring 2012	0.014
GH11_OFU43	spring 2012	0.031

Table 7. List of identified indicator OFUs (after 999 permutations) featuring ≥ 0.75 values for both sensitivity and specificity and P-value < 0.5.

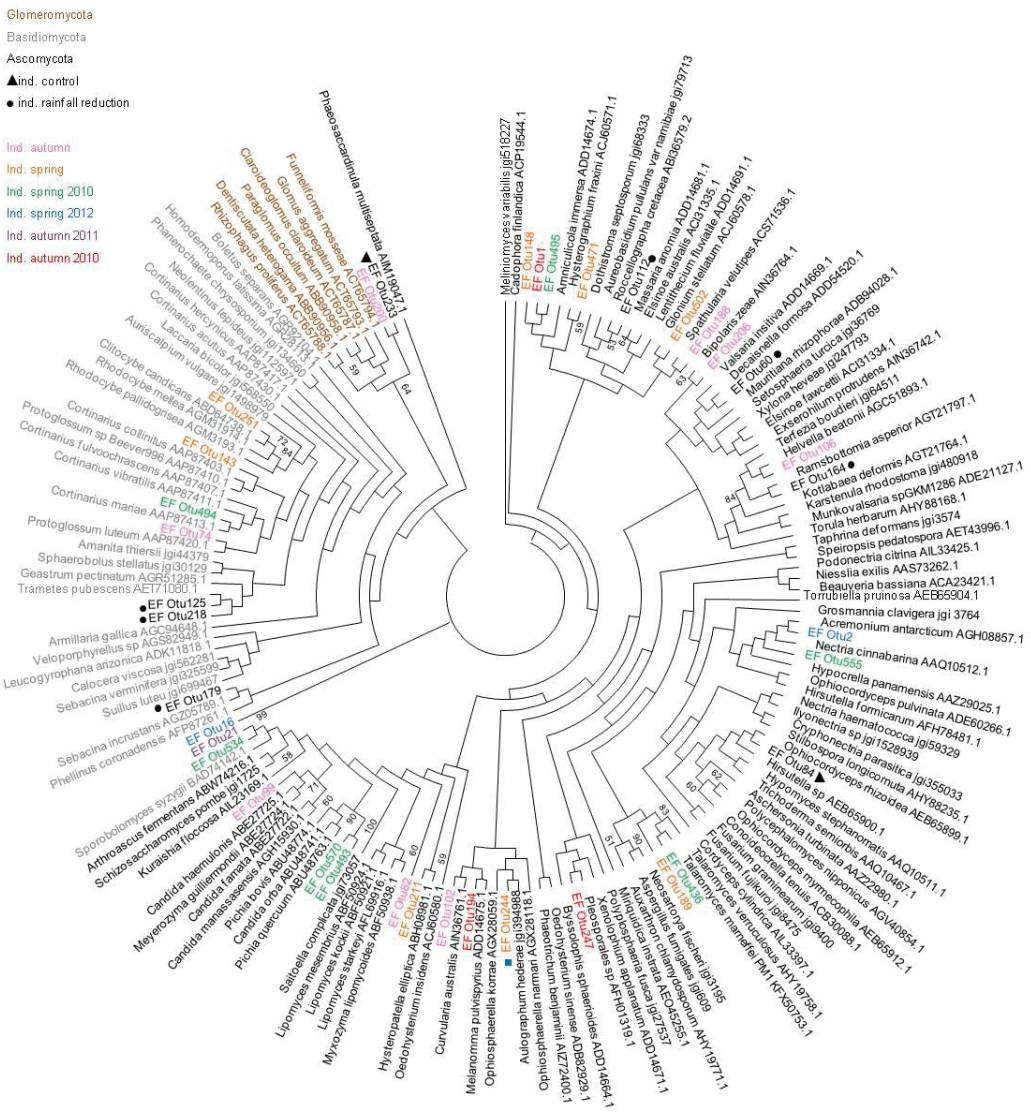

Figure 6. Maximum-likelihood phylogenetic tree of EF1 alpha protein sequences. Indicator EF1-α sequences together with their most similar amino-acid sequences retrieved from public databases were used for building the tree. Only boostrap values ≥50% (100 replicates) are reported. Only the tree topology is represented; lengths of the branches are unrelated to rates of amino acid changes per site.

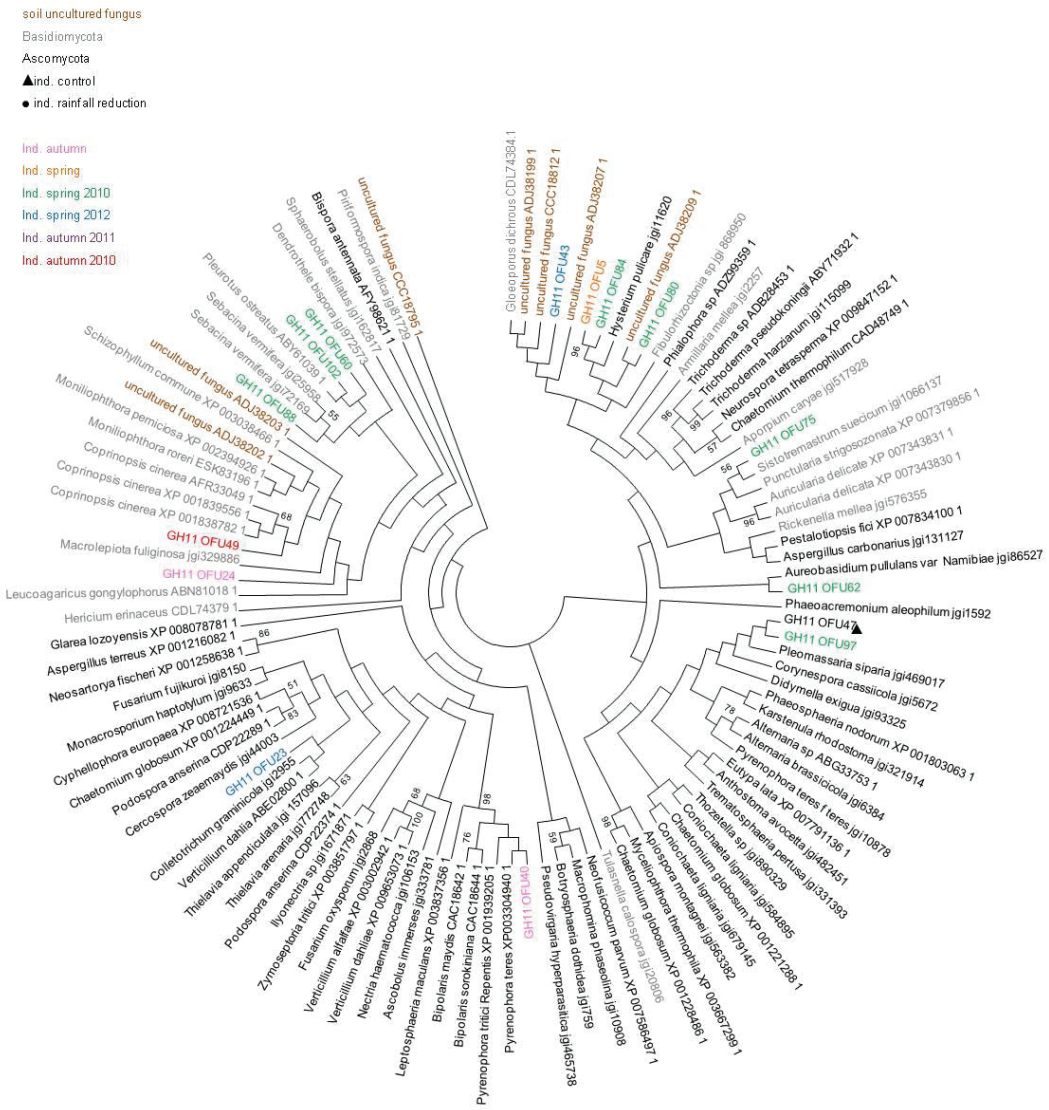

Figure 7. Maximum-likelihood phylogenetic tree of GH11 amino acid sequences. Indicator GH11 sequences together with their most similar amino-acid sequences retrieved from public databases were used for building the tree. Only boostrap values ≥50 (100 replicates) are reported. Only the tree topology is represented; lengths of the branches are unrelated to rates of amino acid changes per site.

3.4.6 *Correlations between communities of different genes*

We investigated whether the different genes displayed similar patterns of β diversity, as revealed independently for each gene by computing either Jaccard, GuniFrac or Bray-Curtis distances between all pairs of samples. This was tested by running Mantel tests on all pairs of "gene-specific" matrices for all three estimators of ß diversity.

Significant positive correlations between phylogenetic distances (GUniFrac), Bray-Curtis indices and Jaccard distances were indeed observed between most genes (Table 8; Fig. S6). The least significant correlations (P>0.014) and the non-significant ones (P>0.05) almost always concerned comparisons with the AA2 gene family (Table 8; Fig. S6).

GuniFrac distance matrix1	***GuniFrac distance matrix 2***	***r***	***P-value***
Fungal EF1-α	GH11	0.3686	**0.001**
Fungal EF1-α	GH7	0.3187	**0.015**
Fungal EF1-α	AA2	0.2096	0.077
GH11	GH7	0.3316	**0.001**
GH11	AA2	0.2125	**0.031**
GH7	AA2	0.1197	0.196
Bray-Curtis distance matrix1	***Bray-Curtis distance matrix 2***	***r***	***P-value***
Fungal EF1-α	GH11	0.3893	**0.001**
Fungal EF1-α	GH7	0.4217	**0.003**
Fungal EF1-α	AA2	0.2627	**0.041**
GH11	GH7	0.3133	**0.001**
GH11	AA2	0.2239	**0.029**
GH7	AA2	0.2052	0.091
Jaccard distance matrix1	***Jaccard distance matrix 2***	***r***	***P-value***
Fungal EF1-α	GH11	0.6214	**0.002**
Fungal EF1-α	GH7	0.4579	**0.005**
Fungal EF1-α	AA2	0.3545	**0.047**
GH11	GH7	0.5349	**0.001**
GH11	AA2	0.3586	**0.014**
GH7	AA2	0.2083	0.081

Table 8. Mantel test results (999 permutations) between GuniFrac distance matrices, Bray-Curtis distance matrices and Jaccard distance matrices, *r:* Mantel statistic, significant P-values (<0.05) are in bold type.

3.5 Discussion

This study represents a first attempt to address the impact of climate change on the diversity of soil fungal communities by means of a parallel high-throughput metabarcoding, on soil-extracted RNA, of four functional genes. One of the target genes (EF1-alpha) is representative of the global taxonomic diversity of the fungal guilds, whereas the remaining three genes are involved in a key process largely controlled by soil fungi, as they encode enzymes active on three different components of the lignocellulosic plant biomass (cellulose, hemicellulose and lignin). Thus far, similar studies have been either restricted to a single taxonomically-informative maker gene (usually the ITS sequence) amplified from soil DNA and/or to a single "functional marker gene" (usually cellulose-encoding GH7 sequences or laccases) from either soil-extracted DNA (Weber *et al.* 2011, Mueller *et al.* 2014) or RNA (Weber *et al.* 2012, Baldrian *et al.* 2012, Kellner *et al.* 2014).

Our study addressed the impact of rainfall reduction on a Mediterranean forest and was carried out in the forest of Puéchabon, a long-term environmental research (LTER) observatory, in which a replicated and continuous throughfall reduction had been implemented for seven years at the start of the present study. The throughfall reduction experiment reduced the amount of leaf litter fall (Limousin *et al.* 2009), the trees in the treatment plot allocate less C to the leaves (reduction in the LAI) and potentially also to the roots.

Three main conclusions can be drawn from the analysis of our sequence datasets. The first one is that none of the different studied parameters, not only rainfall reduction, but also sampling time (date *per se* or season), significantly impacted alpha-diversity estimators (e.g. richness and Shannon indices) for the four different target genes. The second conclusion is that the implemented rainfall reduction seemingly also did not significantly impact the different beta-diversity estimators (Bray-Curtis, Jaccard and GUnifrac distances). Exceptions to this second conclusion concern Jaccard indices for the GH11 and EF1-alpha sequences. As this latter index is based on sequence type (OFU) presence/absence, this could imply that the pools of rare GH11 and EF1-alpha OFUs significantly differ between control plots and plots subjected to rainfall reduction. At first sight, these first and second conclusions could suggest that at Puéchabon, soil fungal communities are fairly stable components of the ecosystem, resistant to both temporal and/or environmental changes. This is however not the case as our third conclusion is that sampling time (and to a lesser extent season) has a strong impact on all three computed beta-diversity indices for the four studied genes as indicated by PERMANOVA analyses.

Our results contrast with the results reported by Richard *et al.* (2011) for the same Puéchabon site between 2007 and 2009 for the Ectomycorrhizal (ECM) fungal root tips community. They observed that five consecutive years of throughfall

reduction indeed induced significant shifts in the community composition and seasonal dynamics of ECM assemblages but did not cause any decrease in species richness or diversity.

However, an absence of measurable effects of rainfall reduction on the diversity of the studied genes is nevertheless in line with the results reported by other authors. Cregger *et al.* (2012) experimentally manipulated precipitation intensity in a semiarid piñon-juniper woodland and assessed soil bacterial and fungal communities diversity and structure by T-RFLP profiling, demonstrating a greater role of seasonal variability in precipitation in determining the composition of soil microbial communities compared to the effect of the experimental precipitation treatments. The same tendency was observed in two Mediterranean ecosystems (Prades holm-oak forest and Garraf schrubland) in Spain by Curiel-Yuste *et al.* (2011, 2014). Jumponnen & Jones (2014), who investigated the effect of a manipulated altered precipitation regime on the soil fungal community in a tallgrass prairie, also reported the same result, suggesting compositional resilience of these communities. As this latter study was based on the sequencing of ITS amplified from soil DNA, the authors highlighted the need of additional studies to address the functional responses of the communities to these environmental manipulations.

Several hypotheses can be put forwards to explain our observations. Firstly, the rainfall reduction treatment, although significantly restricting plant water availability as measured by a decrease in the leaf surface index (Limousin *et al.* 2009), could in itself not be perceived by soil microbial communities. A second hypothesis is that microbial communities present in environments which naturally experience strong and recurrent climatic variations have developed adaptive strategies to cope to these variations and may be resistant/resilient, to some extent, to further increases in their amplitudes. This is particularly relevant for water availability in Mediterranean areas, which are naturally characterized by strong seasonal variations in the levels of precipitations. This hypothesis has recently gained experimental support in several studies which compared the responses to stress of communities which had been or not subjected to a similar stress in the past (Lau & Lennon 2012, Li *et al.* 2014, Thion & Prosser 2014).

The second hypothesis also suggests that the observed significant temporal variations in community structure which affect all four studied genes may not be random but rather repetitive and follow natural climatic cycles and/or the associated ecosystem processes (as illustrated in Fig. 2 for NEE). Similar temporal variations have been reported in forest ecosystems for root associated ectomycorrhizal symbionts and their enzyme activities (Courty *et al.* 2007).

A rigorous test of this hypothesis would require the continuous monitoring of the structure of fungal communities sampled at a fine time scale (e.g. weekly) over one or several years, along with the measurements of several biotic (e.g. tree

physiology) and abiotic (climatic variables) parameters. Only such a survey could disentangle the immediate effects of short term environmental variations (e.g. rewetting of a dry soil) from more long term cyclic seasonal variations on community structure. It is indeed noteworthy noting that altogether the different factors we considered in the present study contributed to at most 25% of the total variation observed in community structure, indicating that many other uncontrolled variables contributed to the observed variation.

One of the current limitation of metabarcoding fungal communities using enzyme-coding genes is that it is currently not possible to assign most, if not all, sequences to a fungal taxon with a high resolution (not only at the species level but also at the genus, family, order and even higher taxonomic levels). Regarding lignocellulolytic genes, this difficulty not only results from an insufficient number of reference sequences in databases but is also exacerbated by the complex evolutionary histories of the corresponding gene families, often represented by different and distantly related paralogous copies in a single fungal genome. The impossibility of taxonomically assigning environmental sequences is particularly relevant in a context of parallel metabarcoding microbial communities using different marker genes. It indeed prevented us to associate to a shared taxon or separate taxa indicator OFUs associated to the same environmental condition. This is not only true for OFUs belonging to different gene families (e.g. GH11 and EF1-alpha), but also to the same family, as in the case of the GH11 one.

We observed that beta diversity indices (Jaccard, Bray-Curtis and UniFrac distances) computed for different gene families were, in most cases, correlated to each other, with the exception of comparisons including the AA2 (basidiomycete class II peroxidase) gene family. One of the simplest explanations accounting for these correlations is that EF1alpha, GH11 and GH7 OFUs originate from the same pool of fungal species expressing each of these genes in a specific soil sample. We know that these different gene families originate from "nested" pools of fungal species. The EF1-alpha gene is a housekeeping gene present in one or several copies per genome in all species whatever their taxonomic identity and trophic strategies. GH7 and GH11 genes often co-occur in variable copy numbers in the genomes of usually saprotrophic and plant pathogenic Ascomycota and Basidiomycota species (e.g. Floudas *et al.* 2012, Barbi *et al.* 2014). As for the AA2 genes, they are rare among Ascomycota species and furthermore the PCR primers we used were specific for Basidiomycota sequences, which are not uniformly distributed in this phylum. The lack of significant correlation between AA2 beta diversity indices and indices calculated for the other genes could therefore be explained by an incongruence between the taxonomic distribution of AA2 genes and the taxonomic distribution of the other genes. Alternatively, it could also be pointed out that levels of soil oxidase activities (partially explained by peroxidase activities) were found to be

disconnected from levels of hydrolytic activities (as performed by GH7 and GH11 gene products) (Talbot *et al.* 2014). We could hypothesize that this functional observation generated the lack of correlation between AA2 beta diversity indices and similar indices computed for the GH7 and GH11 sequences encoding hydrolytic enzymes.

In conclusion, the communities of active fungi and the assemblages of expressed fungal functional genes involved in POM degradation are temporally dynamic in the Mediterranean ecosystem under study. Since the passage of time is associated with natural strong fluctuations in climatic conditions in the Mediterranean area (including the succession of long drought periods during hot summers and cool wet winters), the natural dynamics of soil fungi may be linked to their ability to cope with further alterations in precipitation, such as those expected as a consequence of climate change in this biome.

3.6 Acknowledgements

We thank Audrey Dubost and Stefano Ghignone for their help with bioinformatics tools. This work was supported by ANR project Eumetasol 09-GENM-033-001, CNRS-INSU ECCO Microbien program, EcoFINDERS (FP7-264465) and by Region Rhône–Alpes grant CMIRA 2013 Accueil DOC.

3.7 References

Allard, V., Ourcival, J.M., Rambal, S., Joffre, R., Rocheteau, A. (2008). Seasonal and annual variation of carbon exchange in an evergreen Mediterranean forest in southern France. *Glob. Chang. Biol*. 14:714–725.

Allison, S.D. and Martiny, J.B.H. (2008). Colloquium paper: resistance, resilience, and redundancy in microbial communities. *Proc. Natl. Acad. Sci. U. S. A*. 105 Suppl:11512–9.

Aponte, C., García, L. V., Pérez-Ramos, I.M., Gutiérrez, E., Marañón, T. (2011). Oak trees and soil interactions in Mediterranean forests: a positive feedback model. *J. Veg. Sci.* 22:856–867.

Baldrian, P., Merhautová, V., Petránková, M., Cajthaml, T., Šnajdr, J. (2010). Distribution of microbial biomass and activity of extracellular enzymes in a hardwood forest soil reflect soil moisture content. *Appl. Soil Ecol.* 46:177–182.

Baldrian, P., Kolařík, M., Stursová, M., Kopecký, J., Valášková, V., Větrovský, T., *et al.* (2012). Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. *ISME J.* 6:248–58.

Baldrian P, Šnajdr J, Merhautová V, Dobiášová P, Cajthaml T, Valášková V. (2013). Responses of the extracellular enzyme activities in hardwood forest to soil temperature and seasonality and the potential effects of climate change. Soil Biol. Biochem. 56:60–68.

Barbi, F., Bragalini, C., Vallon, L., Prudent, E., Dubost, A., Fraissinet-Tachet, L., *et al.* (2014). PCR primers to study the diversity of expressed fungal genes encoding lignocellulolytic enzymes in soils using high-throughput sequencing. *PLoS One* 9:e116264.

Bell, C.W., Acosta-Martinez, V., McIntyre, N.E., Cox, S., Tissue, D.T., Zak, J.C. (2009). Linking microbial community structure and function to seasonal differences in soil moisture and temperature in a Chihuahuan desert grassland. *Microb. Ecol.* 58:827–42.

Berg, B., McClaugherty, C. (2008). Plant Litter. Decomposition, humus formation, carbon sequestration, 2nd edition, Springer-Verlag, Berlin, Heidelberg

Bödeker, I.T.M., Nygren, C.M.R., Taylor, A.F.S., Olson, A., Lindahl, B.D. (2009). ClassII peroxidase-encoding genes are present in a phylogenetically wide range of ectomycorrhizal fungi. *ISME J.* 3:1387–95.

De Cáceres, M. and Legendre, P. (2009). Associations between species and groups of sites: indices and statistical inference. *Ecology* 90:3566–3574.

Castro, H.F., Classen, A.T., Austin, E.E., Norby, R.J., Schadt, C.W. (2010). Soil microbial community responses to multiple experimental climate change drivers. *Appl. Environ. Microbiol.* 76:999–1007.

Chen, J., Bittinger, K., Charlson, E.S., Hoffmann, C., Lewis, J., Wu, G.D., *et al.* (2012). Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics* 28:2106–13.

Courty, P.E., Bréda, N., Garbaye, J. (2007). Relation between oak tree phenology and the secretion of organic matter degrading enzymes by *Lactarius quietus* ectomycorrhizas before and during bud break. *Soil Biol. Biochem*. 39:1655–1663

Cregger ,M.A., Schadt, C.W, McDowell, N.G., Pockman, W.T, Classen, A.T. (2012). Response of the soil microbial community to changes in precipitation in a semiarid ecosystem. *Appl. Environ. Microbiol.* 78:8587–94.

Cruz-Martínez, K., Suttle, K.B., Brodie, E.L., Power, M.E., Andersen, G.L., Banfield, J.F. (2009). Despite strong seasonal responses, soil microbial consortia are more resilient to long-term changes in rainfall than overlying grassland. *ISME J.* 3:738–44.

Curiel Yuste, J., Peñuelas, J., Estiarte, M., Garcia-Mas, J., Mattana, S., Ogaya, R., *et al.* (2011). Drought-resistant fungi control soil organic matter decomposition and its response to temperature. *Glob. Chang. Biol.* 17:1475–1486.

Curiel Yuste, J., Fernandez-Gonzalez, A.J., Fernandez-Lopez, M., Ogaya, R., Penuelas, J., Sardans, J., Lloret, F. (2014). Strong functional stability of soil microbial communities under semiarid Mediterranean conditions and subjected to long-term shifts in baseline precipitation. *Soil Biol. Biochem. 69: 223–233.*

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–7.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27:2194–200.

Edwards, I.P., Upchurch, R. A., Zak, D.R. (2008). Isolation of fungal cellobiohydrolase I genes from sporocarps and forest soils by PCR. *Appl. Environ. Microbiol*. 74:3481–9.

Floudas, D., Binder, M., Riley, R., Barry, K., Blanchette, R. a, Henrissat, B., *et al.* (2012). The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science 336*(6089), 1715–9.

García-Palacios, P., Vandegehuchte, M.L., Ashley, S. E., Dam, M., Post, K.H., Ramirez, K.S., *et al.* (2014). Are there links between responses of soil microbes and ecosystem functioning to elevated CO2 , N deposition and warming? A global perspective. *Glob. Chang. Biol.* DOI:10.1111/gcb.12788

Gibelin, A. and Déqué, M. (2003). Anthropogenic climate change over the Mediterranean region simulated by a global variable resolution model. *Clim. Dyn.* 20:327–339.

Gihring, T.M., Green, S.J., Schadt, C.W. (2012). Massively parallel rRNA gene sequencing exacerbates the potential for biased community diversity comparisons due to variable library sizes. *Environ. Microbiol*. 14:285–90

Giorgi, F. and Lionello, P. (2008). Climate change projections for the Mediterranean region. *Glob. Planet. Change* 63:90–104.
Gouy, M., Guindon, S., Gascuel, O. (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* 27:221–4.
Hawkes, C. V., Kivlin, S.N., Rocca, J.D., Huguet, V., Thomsen, M. A., Suttle, K.B. (2011). Fungal community responses to precipitation. *Glob. Chang. Biol.* 17:1637–1645.
IPCC, Intergovernmental panel on climate change. (2007). Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change.
Jumpponen, A. And Jones, K L., (2014). Tallgrass prairie soil fungal communities are resilient to climate change. *Fungal Ecology* (10):44–57
Kaiser, C., Koranda, M., Kitzler, B., Fuchslueger, L., Schnecker, J., Schweiger, P., *et al.* (2010). Belowground carbon allocation by trees drives seasonal patterns of extracellular enzyme activities by altering microbial community composition in a beech forest soil. *New Phytol* 187:843–58.
Kellner, H., Vandenbol, M. (2010). Fungi Unearthed: Transcripts Encoding Lignocellulolytic and Chitinolytic Enzymes in Forest Soil. *PLoS ONE* 5(6): e10971.
Kellner, H., Luis, P., Pecyna, M.J., Barbi, F., Kapturska, D., Krüger, D., *et al.* (2014). Widespread occurrence of expressed fungal secretory peroxidases in forest soils. *PLoS ONE* 9:e95557.
Kandeler, E., Stemmer, M., & Gerzabek, M. H. (2005). Role of Microorganisms in Carbon Cycling in Soils. In A. Varma & B. Francois (Eds.), *Microorganisms in Soils: Roles in Genesis and Functions* (pp. 139–157). Springer Berlin Heidelberg.
Landesman, W.J., Dighton, J. (2011). Shifts in microbial biomass and the bacteria: fungi ratio occur under field conditions within 3 h after rainfall. *Microb. Ecol.* 62:228–36.
Lau, J.A., Lennon, J.T, (2012). Rapid responses of soil microorganisms improve plant fiteness in novel environments. *Proc. Natl. Acad. Sci. U. S. A.* 109(35):14058-62.
Li, J.,Zheng, Y.M, Liu, Y.R,Ma, Y.B.,Hu, H.W, He, J.Z. (2014). Initial copper stress strengthens the resistance of soil microorganisms to a subsequent copper stress. *Microbial Ecology* 67(4):931-941.
Limousin, J., Rambal, S., Ourcival ,J., Joffre, R. (2008). Modelling rainfall interception in a mediterranean Quercus ilex ecosystem: Lesson from a throughfall exclusion experiment. *J. Hydrol.* 357:57–66.
Limousin, J Rambal, S., Ourcival J., Rocheteau, A., Joffre, R., Rodriguez- Cortina, R., (2009), Long-term transpiration change with rainfall decline in a Mediterranean *Quercus ilex* forest. *Global Change Biology* 15, 2163–75
Lindahl, B. D., & Tunlid, A. (2014). Tansley insight Ectomycorrhizal fungi – potential organic matter decomposers , yet not saprotrophs. *New Phytol.* doi:10.1111/nph.13201
Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42:D490–5.
Luis, P., Kellner,H., Martin, F., Buscot, F. (2005). A molecular method to evaluate basidiomycete laccase gene expression in forest soils. *Geoderma* 128:18–27.
Luis, P., Walther, G., Kellner, H., Martin, F., Buscot, F. (2004). Diversity of laccase genes from basidiomycetes in a forest soil. *Soil Biol. Biochem.* 36:1025–1036.
Maijala, P., Harrington, T.C., Raudaskoski, M. (2003). A peroxidase gene family and gene trees in Heterobasidion and related genera. Mycologia 95:209–221.
McMurdie, P.J. and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8:e61217.
Mouillot, F., Rambal, S., Joffre, R. (2002). Simulating climate change impacts on fire frequency and vegetation dynamics in a Mediterranean-type ecosystem. *Glob. Chang. Biol*. 423–437.

Mueller, R.C., Balasch, M.M., Kuske, C.R. (2014). Contrasting soil fungal community responses to experimental nitrogen addition using the large subunit rRNA taxonomic marker and cellobiohydrolase I functional marker. *Mol. Ecol*. 23:4406-44017

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L. , Solymos, P., Stevens, M. H. H., Wagner. H.(2013). vegan: Community Ecology Package. http://cran.r-project.org/package=vegan.

R Core Team. (2014). R: A language and environment for statistical computing. http://www.R-project.org/.

Rehner, S. and Buckley, E. (2005). A Beauveria phylogeny inferred from nuclear ITS and EF1-α sequences: evidence for cryptic diversification and links to Cordyceps teleomorphs. *Mycologia* 97:84–98.

Richard, F., Roy, M., Shahin, O., Sthultz, C., Duchemin, M.,Joffre, R., Selosse, M. (2011) Ectomycorrhizal communities in a Mediterranean forest ecosystem dominated by Quercus ilex: seasonal dynamics and response to drought in the surface organic horizon. *Annals of Forest Science* 68(1):57-68.

Rineau, F., Shah, F., Smits, M.M., Persson, P., Johansson, T., Carleer, R., *et al.* (2013). Carbon availability triggers the decomposition of plant litter and assimilation of nitrogen by an ectomycorrhizal fungus. *ISME J.* 7:2010–22.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., *et al.* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol*. 75:7537–41.

Schmitt, A. and Glaser, B. (2011). Organic matter dynamics in a temperate forest as influenced by soil frost. *J. Plant Nutr. Soil Sci.* 174:754–764.

Schneider, T., Keiblinger, K.M., Schmid, E., Sterflinger-Gleixner, K., Ellersdorfer, G., Roschitzki,B., *et al.* (2012). Who is who in litter decomposition? Metaproteomics reveals major microbial players and their biogeochemical functions. *ISME J.* 6:1749–62.

Talbot, J.M., Bruns, T.D., Smith, D.P., Branco, S., Glassman, S.I,. Erlandson, S., *et al.* (2013). Independent roles of ectomycorrhizal and saprotrophic communities in soil organic matter decomposition. *Soil Biol. Biochem.* 57:282–291.

Talbot, J.M., Bruns, T.D., Taylor, J.W., Smith, D.P., Branco, S., Glassman, S.I., *et al.* (2014). Endemism and functional convergence across the North American soil mycobiome. *Proc. Natl. Acad. Sci. U. S. A* 111(17):6341-46

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30:2725–9.

Thion, C ., Prosser, J.I. (2014). Differential response of nonadapted ammonia-oxidising archaea and bacteria to drying-rewetting stress.*Fems Microbiology Ecology* 90(2):380-389.

Toberman, H., Freeman, C., Evans, C., Fenner, N., Artz, R.R.E. (2008). Summer drought decreases soil fungal diversity and associated phenol oxidase activity in upland Calluna heathland soil. *FEMS Microbiol. Ecol.* 66:426–36.

Trivedi, P., Anderson, I.C., Singh, B.K. (2013). Microbial modulators of soil carbon storage: integrating genomic and metabolic knowledge for global prediction. *Trends Microbiol.* 21:641–51.

Wagg, C., Bender, S.F., Widmer, F., van der Heijden, M.G. A. (2014). Soil biodiversity and soil community composition determine ecosystem multifunctionality. *Proc. Natl. Acad. Sci. U. S. A.* 111:5266–70.

Wang, Q., Quensen 3rd, J.F., Fish, J.A., Lee, T.K., Sun, Y., Tiedje, J.M., *et al.* (2013). Ecological patterns of *nifH* genes in four terrestrial climatic zones explored with targeted metagenomics using FrameBot, a new informatics tool. *MBio* 4:e00592–13.

Weber, C.F., Zak, D.R., Hungate B. A., Jackson, R.B., Vilgalys, R., Evans, R.D., *et al.* (2011). Responses of soil cellulolytic fungal communities to elevated atmospheric CO_2 are complex and variable across five ecosystems. *Environ. Microbiol.* 13:2778–93.

Weber, C.F., Balasch, M.M., Gossage, Z., Porras-Alfaro, A., Kuske, C.R. (2012). Soil fungal cellobiohydrolase I gene (*cbhI*) composition and expression in a loblolly pine plantation under conditions of elevated atmospheric CO2 and nitrogen fertilization. *Appl. Environ. Microbiol.* 78:3950–7.

Zeglin, L., Bottomley, P., Jumpponen A. (2013). Altered precipitation regime affects the function and composition of soil microbial communities on multiple time scales. *Ecology* 94:2334–2345.

Zhang, J., Kobert, K., Flouri, T., Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30:614–20.

3.8 Supplementary data

Sample Name	*plot*	*Treatment*	*Sample collection date and time*	*Latitude of sample collection*	*Longitude of sample collection*	*Altitude (m.s.l)*	*Soil water content (%)*
P10 MR R1	1	control	21/04/2010 (10-12h)	N43.74155	E3.59697	259	33.92
P10 MR R2	2	control	21/04/2010 (10-12h)	N43.74083	E3.59535	263	29.26
P10 MD R1	1	rainfall reduction	21/04/2010 (10-12h)	N43.74163	E3.59678	260	32.41
P10 MD R2	2	rainfall reduction	21/04/2010 (10-12h)	N43.74076	E3.59554	262	25.79
A10 MR R1	1	control	03/12/2010 (10-12h)	N43.74155	E3.59697	259	40.2
A10 MR R2	2	control	03/12/2010 (10-12h)	N43.74155	E3.59697	259	40.73
A10 MR R3	3	control	03/12/2010 (10-12h)	N43.74158	E3.59602	263	34.72
A10 MD R1	1	rainfall reduction	03/12/2010 (10-12h)	N43.74163	E3.59678	260	39.2
A10 MD R2	2	rainfall reduction	03/12/2010 (10-12h)	N43.74076	E3.59554	262	38.33
A10 MD R3	3	rainfall reduction	03/12/2010 (10-12h)	N43.74148	E3.59624	259	35.93
A11 MR R1	1	control	02/12/11 (10-12h)	N43.74155	E3.59697	259	32.7
A11 MR R2	2	control	02/12/11 (10-12h)	N43.74155	E3.59697	259	28.77
A11 MR R3	3	control	02/12/11 (10-12h)	N43.74158	E3.59602	263	22.95
A11 MD R1	1	rainfall reduction	02/12/11 (10-12h)	N43.74163	E3.59678	260	29.22
A11 MD R2	2	rainfall reduction	02/12/11 (10-12h)	N43.74076	E3.59554	262	23.44
A11 MD R3	3	rainfall reduction	02/12/11 (10-12h)	N43.74148	E3.59624	259	27.93
P12 MR R1	1	control	16/05/2012 (10-12h)	N43.74155	E3.59697	259	27.18
P12 MR R2	2	control	16/05/2012 (10-12h)	N43.74155	E3.59697	259	21.25
P12 MR R3	3	control	16/05/2012 (10-12h)	N43.74158	E3.59602	263	22.62
P12 MD R1	1	rainfall reduction	16/05/2012 (10-12h)	N43.74163	E3.59678	260	20.95
P12 MD R2	2	rainfall reduction	16/05/2012 (10-12h)	N43.74076	E3.59554	262	15.19
P12 MD R3	3	rainfall reduction	16/05/2012 (10-12h)	N43.74148	E3.59624	259	17.38

Table S1. List of the soil samples used in the present study.

	control	***rainfall reduction***
Total Nitrogen content (g/kg)	6.12	6.43
Organic Carbon (g/kg)	122	115
C/N	19.9	17.9
Organic matter (g/kg)	211	199
pH H20	7.2	7.36
pH KCl	6.82	7.04
Phosphorous (P2O5, g/kg)	0.021	0.028
CEC cobaltihexamine(cmol+/kg)	43.3	45.7
Aluminium (Al,cmol+/kg)	0.0965	0.0879
Calcium (Ca, cmol+/kg)	43.7	46
Iron (Fe, cmol+/kg)	0.0187	0.0183
Magnesium (Mg, cmol+/kg)	2.2	2.25
Manganese (Mn, cmol+/kg)	0.0636	0.0457
Potassium (K, cmol+/kg)	1	1.13
Sodium (Na, cmol+/kg)	0.136	0.125

Table S2. Characteristics of soils sampled in control and rainfall reduction treatment plots.

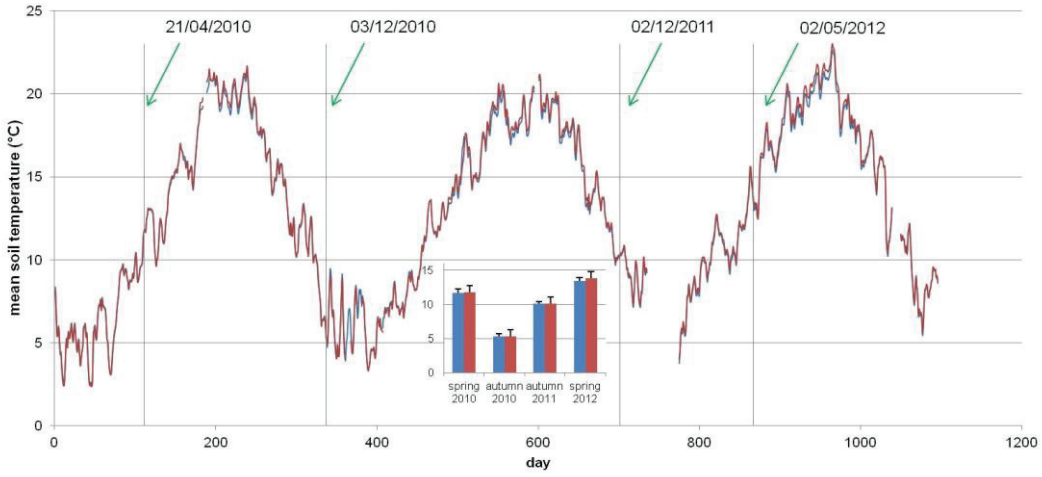

Figure S1 . Variations in mean soil temperature (°C) of the control (blue line) and rainfall reduction (red line) plots (from January 2010 to December 2012). The four indicated dates correspond to the soil sampling date. Insert: values of the four sampling date.

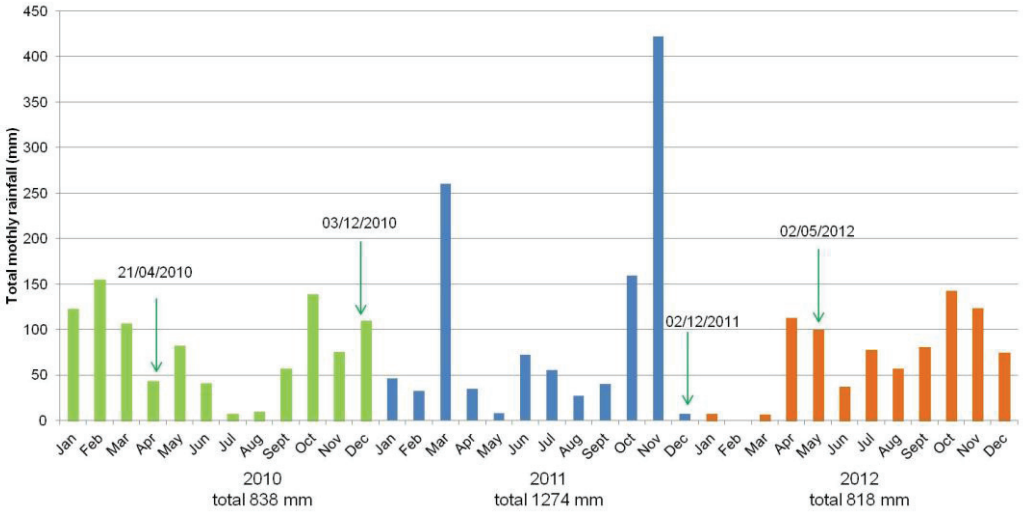

Figure S2 . Total monthly rainfall (mm) measured at the Puéchabon experimental forest in the period 2010-2012. For each year total annual rainfall is reported. Dates/arrows on the graph indicate the days when soil samples were collected.

(A) (B)

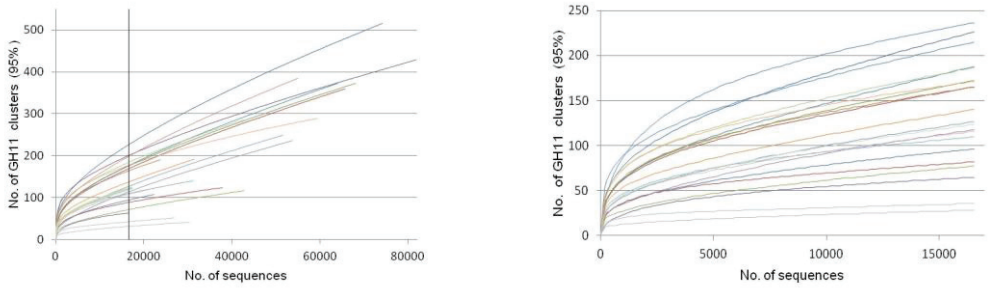

Figure S3 *part1*. Rarefaction curves depicting the effect of sampling depth on the number of OFUs (clusters) identified for each genes (A) before and (B) after sub-sampling to the a 16,554; 4705; 344 and 826 sequence threshold (black vertical lines) for the GH11, Fungal EF1- α, GH7 and AA2 gene families respectively.

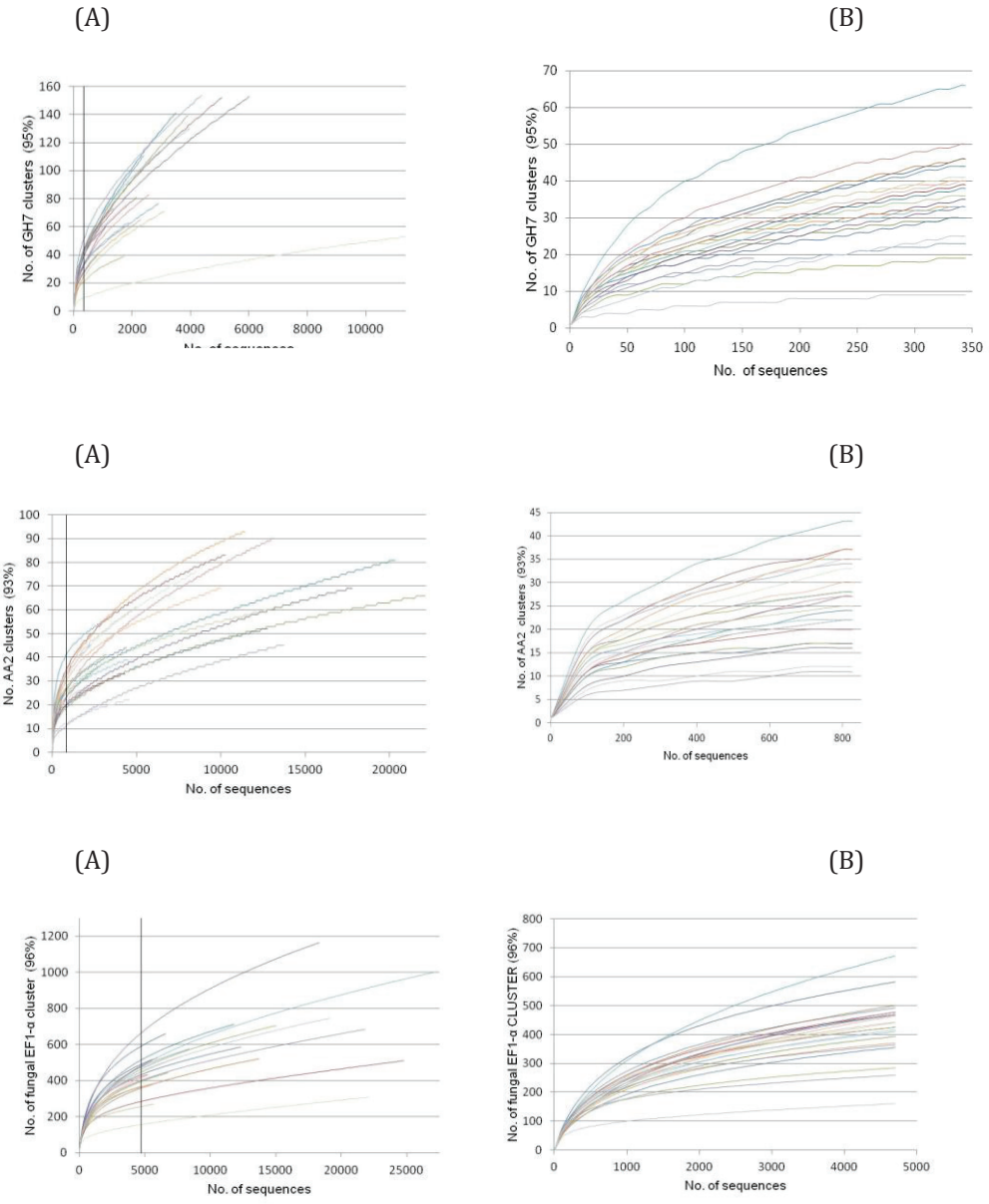

Figure S3 *part2*. Rarefaction curves depicting the effect of sampling depth on the number of OFUs (clusters) identified for each genes (A) before and (B) after sub-sampling to the a 16,554; 4705; 344 and 826 sequence threshold (black vertical lines) for the GH11, Fungal EF1- α, GH7 and AA2 gene families respectively.

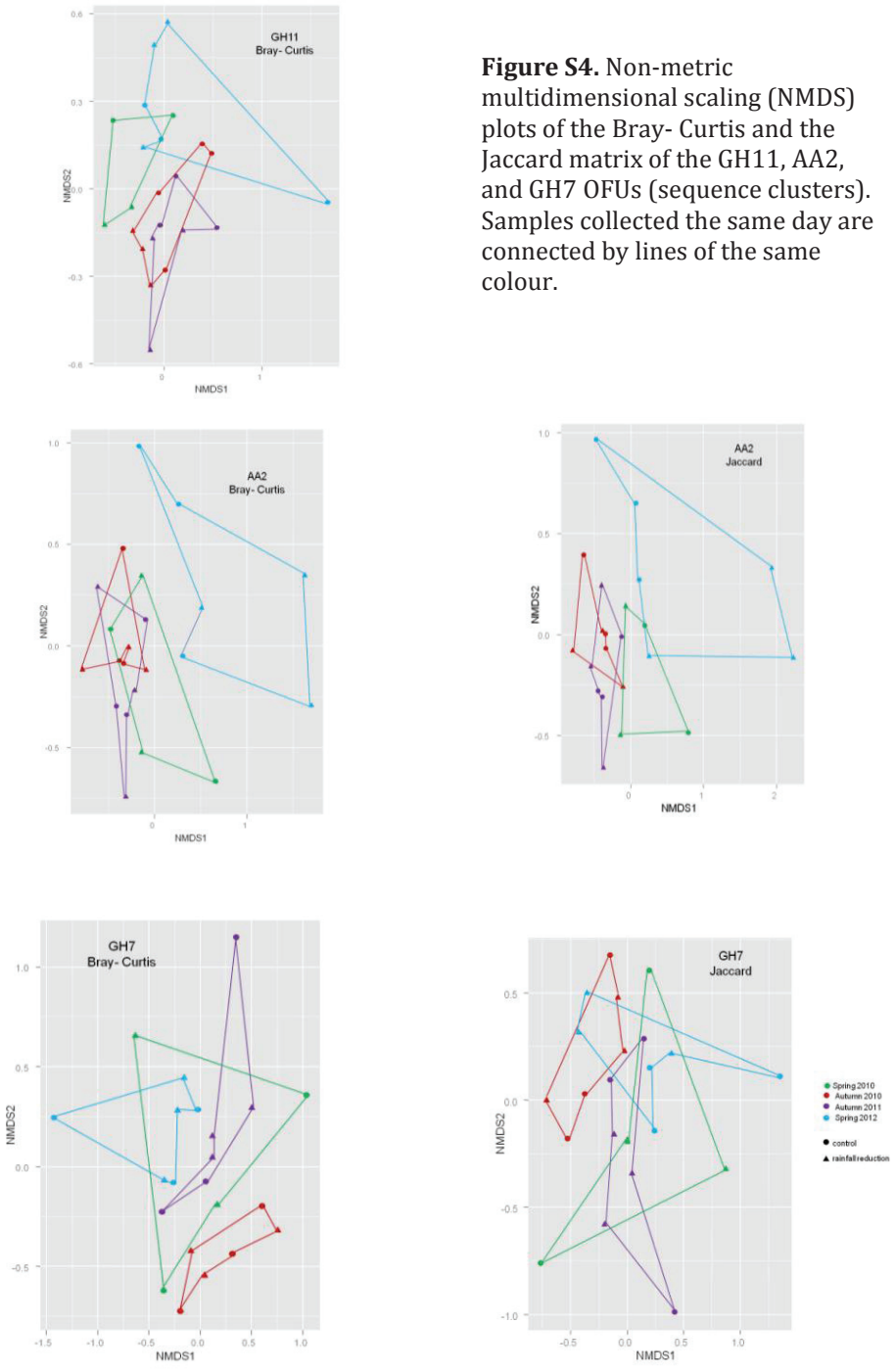

Figure S4. Non-metric multidimensional scaling (NMDS) plots of the Bray- Curtis and the Jaccard matrix of the GH11, AA2, and GH7 OFUs (sequence clusters). Samples collected the same day are connected by lines of the same colour.

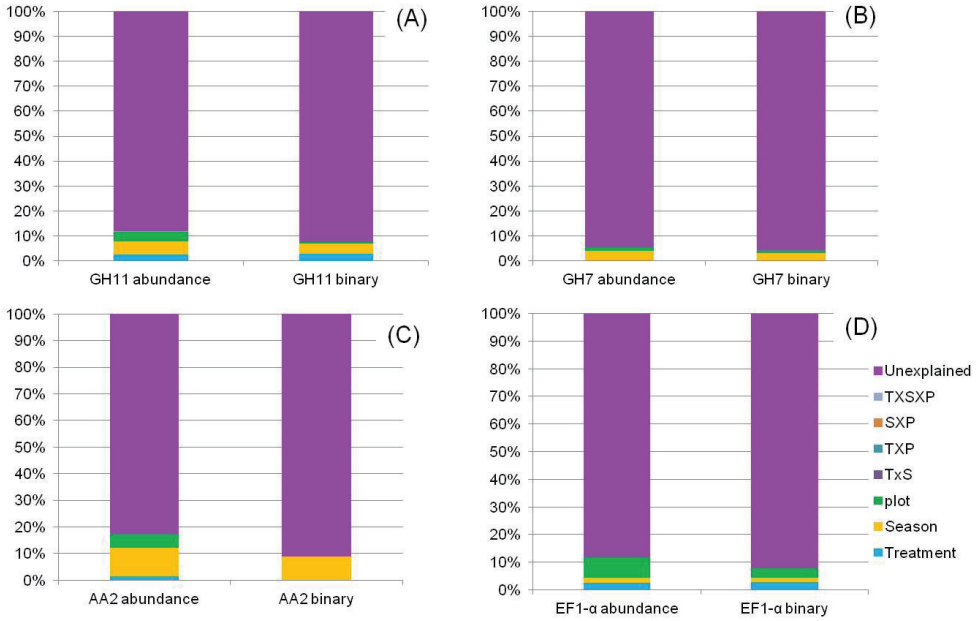

Figure S5. Variance partitioning, displayed as bar plots, giving the relative contribution of treatment, season and plot factors to total community composition. Variance partitioning was calculated using abundance data (after Hellinger transformation) and binary data. (A) GH11, (B) GH7, (C) AA2, (D) Fungal EF1-α community

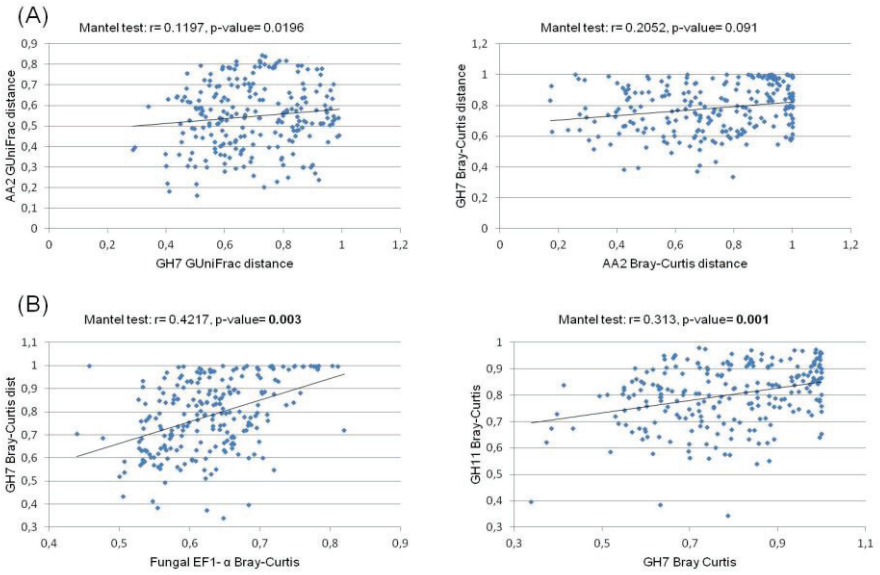

Figure S6. Illustrations of correlations between community dissimilarity matrices computed for different gene families. Pairwise community dissimilarity was calculated for each marker using the GUniFrac or Bray-Curtis distance metric. Examples of (A) non-significant and(B) significant correlations based on Mantel tests (999 permutations) are given.

Chapter IV

Solution hybrid selection capture for the recovery of functional full-length eukaryotic cDNAs from complex environmental samples

4.1 Foreword

The analyses described in the previous chapters were performed to understand the responses of fungal communities to major global change drivers since these organisms play fundamental roles in soil ecosystems. In addition to their ecological importance, thanks to their capacity to secrete powerful enzymes and compounds, they are also major contributors to important both ancient and modern biotechnological processes. Processes and products that utilize fungi include baking, wine making, brewing, and the production of antibiotics, alcohols, enzymes, organic acids, and numerous pharmaceuticals.

Soil fungal communities potentially represent a rich source of natural products including biocatalysts. Culture-independent molecular techniques are powerful tools to discover genes with particular function from uncultured microorganisms in the environment. However, up to now, none of the methods used in environmental genomics allow for a rapid isolation of numerous functional genes especially from eukaryotic microbial communities.

The functional characterization of eukaryotic genes isolated from environmental DNA (metagenomics) faces several problems. The frequent presence of introns make the characterization of functional genes difficult and together with the frequent lack of conservation of motifs in promoter sequences prevent the expression of genomic copies of eukaryotic protein-coding genes not only in a bacterial cell but also in most eukaryotic host.

The use of RNA extracted from environmental samples (metatranscriptomics) could circumvent these two specific problems. However, the total RNA pool in environmental microbial communities consists primarily of ribosomal and transfer RNAs (Karpinets *et al.* 2006), and only about 1-5% messanger RNA (He *et al.* 2010). Owing to their 3' poly-A tails, eukaryotic mRNA can however be specifically isolated from a complex RNA mixture and converted into intron-less cDNAs that, for example, can be cloned to generate environmental metatranscriptomic cDNA libraries.

Many genes transcripts are nevertheless extremely rare. In particular gene transcripts encoding Carbohydrates-Active enZYmes (CAZymes) from forest soil samples represent only less than 1% of the total mRNA pools (Damon *et al.* 2012). In the study presented in the *Chapter III,* in order to characterize a high number of CAZymes transcripts from soil samples we applied a commonly used method based on polymerase chain reaction (PCR) amplification. PCR based methods suffer of primers bias and in addition the use of primers designed on the catalytic domains of the targeted genes, which is often the case in functional genes surveys, does not allow for the recovery of the full-length gene but of only a portion, preventing the functional study of these genes by expression in an heterologous microbial host.

For this reason, the aim of the work described in this chapter was to develop a method to rapidly enrich a cDNAs pool in specific transcripts and to recover full length functional genes from the environment which is a fundamental requirement to study and discover new genes that may have a biotechnological application. Specifically we targeted genes belonging to the CAZymes family Glycoside Hydrolase 11 (GH11), targeted also in the study of the previous chapter, and for which we designed specific primers (*Chapter V*).

GH11 is a family of xylanases which are hemicellulolytic enzymes responsible for the degradation of the heteroxylans present in the lignocellulosic plant cell wall. GH11 is one of the best characterized GH families with bacterial and fungal members and is considered as true xylanases compared to the other families because of their high substrate specificity. GH11 xylanases have for a long time been used as biotechnological tools in various industrial applications. For instance they are extensively employed in food technology (e.g bread making to improve baking properties), in feed technology (in animal feed to increase digestibility) or in the fiber and paper industry (for the biobleaching of kraft pulps) and represent in addition promising candidates for future other uses (Paës *et al.* 2012).

The method was implemented in collaboration with Prof. Pierre Peyret's research group at the University of Auvergne (Clermont Ferrand, France).

References:

Damon, C., Lehembre, F., Oger-Desfeux, C., Luis, P., Ranger, J., Fraissinet-Tachet, L., *et al.* (2012). Metatranscriptomics reveals the diversity of genes expressed by eukaryotes in forest soils. *PLoS One* 7:e28967.

He, S., Wurtzel, O., Singh, K., Froula, J.L., Yilmaz, S., Tringe, S.G., *et al.* (2010). Validation of two ribosomal RNA removal methods for microbial metatranscriptomics. *Nat Meth* 7:807–812.

Karpinets, T.V., Greenwood, D.J., Sams, C.E., Ammons, J.T. (2006). RNA:protein ratio of the unicellular organism as a characteristic of phosphorous and nitrogen stoichiometry and of the cellular requirement of ribosomes for protein synthesis. *BMC Biol.* 4:30.

Paës, G., Berrin, J.G., Beaugrand, J. (2012). GH11 xylanases: Structure/function/properties relationships and applications. *Biotechnol. Adv.* 30:564–92.

DNA Research **21**, 685–694, (2014)
Advance Access publication on 3 October 2014
doi:10.1093/dnares/dsu030

Solution Hybrid Selection Capture for the Recovery of Functional Full-Length Eukaryotic cDNAs From Complex Environmental Samples

Claudia Bragalini[1,2], Céline Ribière[3], Nicolas Parisot[3], Laurent Vallon[2], Elsa Prudent[2], Eric Peyretaillade[3], Mariangela Girlanda[2,4], Pierre Peyret[3], Roland Marmeisse[1,2], and Patricia Luis[2,*]

Department of Life Sciences and Systems Biology, University of Turin, viale Mattioli 25, Turin 10125, Italy[1]; Ecologie Microbienne, UMR CNRS 5557, USC INRA 1364, Université de Lyon, Université Lyon 1, Villeurbanne 69622, France[2]; EA 4678 CIDAM, BP 10448, Clermont Université, Université d'Auvergne, Clermont-Ferrand F-63001, France[3] and Istituto per la Protezione Sostenibile delle Piante (IPSP), Consiglio Nazionale delle Ricerche, Viale Mattioli 25, Turin 10125, Italy[4]

*To whom correspondence should be addressed. Tel. +33 472431050. Fax. +33 47 2431643. Email: patricia.luis@univ-lyon1.fr

Edited by Prof. Takashi Ito
(Received 16 May 2014; accepted 4 September 2014)

Abstract

Eukaryotic microbial communities play key functional roles in soil biology and potentially represent a rich source of natural products including biocatalysts. Culture-independent molecular methods are powerful tools to isolate functional genes from uncultured microorganisms. However, none of the methods used in environmental genomics allow for a rapid isolation of numerous functional genes from eukaryotic microbial communities. We developed an original adaptation of the solution hybrid selection (SHS) for an efficient recovery of functional complementary DNAs (cDNAs) synthesized from soil-extracted polyadenylated mRNAs. This protocol was tested on the Glycoside Hydrolase 11 gene family encoding *endo*-xylanases for which we designed 35 explorative 31-mers capture probes. SHS was implemented on four soil eukaryotic cDNA pools. After two successive rounds of capture, >90% of the resulting cDNAs were GH11 sequences, of which 70% (38 among 53 sequenced genes) were full length. Between 1.5 and 25% of the cloned captured sequences were expressed in *Saccharomyces cerevisiae*. Sequencing of polymerase chain reaction-amplified GH11 gene fragments from the captured sequences highlighted hundreds of phylogenetically diverse sequences that were not yet described, in public databases. This protocol offers the possibility of performing exhaustive exploration of eukaryotic gene families within microbial communities thriving in any type of environment.

Key words: metatranscriptomics; soil RNA; soil eukaryotes; sequence capture; glycoside hydrolase family GH11

1. Introduction

A common objective of many studies in the field of environmental microbiology is to evaluate the functional diversity of the complex microbial communities colonizing natural or man-made environments, fresh or marine waters, sediments, soils, digestive tracts or food products. This diversity can be apprehended through the systematic sequencing and functional annotation of DNA (metagenomics) or RNA (metatranscriptomics) molecules directly extracted from environmental samples.[1,2] However, as a result of the extreme taxonomic richness of most microbial communities, high-throughput shotgun sequencing of environmental nucleic acids is far from covering their full gene repertoire.[3] Alternatively, many studies focus on specific environmental processes which, for some of them, are controlled by a limited and defined set of genes encoding key enzymes. The diversity of the corresponding gene families and of the organisms that possess and

express them is classically evaluated by the systematic sequencing and taxonomic annotation of polymerase chain reaction (PCR)-amplified gene fragments from environmental DNA or RNA (metabarcoding).[4–7] This latter approach has itself well-documented limitations. One of the limitations is that the use of a single pair of degenerate primers, designed to hybridize to internal gene consensus sequences, usually fails to amplify all homologous sequences present in an environmental sample.[8] Another, often underestimated limitation is that metabarcoding does not allow amplification of full-length functional genes. Besides limiting the number of phylogenetically informative nucleotide positions for precise phylogenetic assignment of environmental sequences, obtaining partial sequences also prevents their functional study by expression in a heterologous microbial host. Full-length functional genes are yet of importance (i) in ecology to establish potential relationships between enzyme catalytic properties (substrate range, sensitivity to physicochemical parameters) and prevailing environmental conditions and (ii) in environmental biotechnology to isolate novel biocatalysts for industrial purpose.

Recently, Denonfoux *et al.*[9] developed an alternative strategy to explore microbial communities from complex environments. Based on solution hybrid selection (hereafter referred to as SHS), this method allows for the specific recovery of large DNA fragments harbouring biomarkers of interest even from rare or unknown microorganisms. Indeed, SHS is based on the design of several oligonucleotide probes which can cover the whole gene of interest as opposed to PCR strategies targeting internal regions. Moreover, explorative probe design strategies using appropriate software such as HiSpOD[10] or KASpOD[11] allow recovering not yet described homologous sequences.[9] These probes are synthesized as biotinylated RNA oligonucleotides and hybridized, in solution, to the target gene sequences diluted among a majority of non-target DNA fragments. The hybrid molecules (biotinylated probes + target sequences) are then specifically captured by affinity binding on streptavidin-coated paramagnetic beads. SHS can be repeated several times successively to increase the enrichment in desired sequences by a factor of up to 1.7×10^5 times.[9] In environmental microbiology, the captured DNA fragments can be subjected to high-throughput sequencing. *In silico* assembly of the reads not only leads to the reconstruction of the full-length sequences of the different members of the targeted gene family, but also of their genomic environment and could therefore facilitate operon reconstructions.[9]

In microbial ecology, SHS has thus far been successfully used to capture archaeal protein-coding genes from environmental DNA.[9] As previously discussed,[12] environmental DNA is however not the most appropriate matrix to recover full-length functional genes of eukaryotic origin, which could be easily expressed in a heterologous microbial host. Environmental polyadenylated messenger RNAs, devoid of introns, represent a better source of eukaryotic genes which, following their conversion into complementary DNAs (cDNAs), can be expressed in either bacteria or yeasts.[12–16]

Soil eukaryotes such as fungi are highly diverse,[17,18] play essential roles in soil biology as, for example, the main agents in plant organic matter degradation[19,20] and represent a rich source of enzymes and biomolecules used in industry.[21] Despite these obvious interests, very few environmental genomics studies specifically focus on soil eukaryote functional diversity.[22]

To promote such studies, we developed and evaluated in the present report an original adaptation of the SHS for the efficient recovery of full-length functional fungal cDNAs synthesized from soil RNA. Successful development of this technique was favoured by the ever increasing number of available fungal genomes that provide a correspondingly large number of members of specific gene families for the design of hybridization probes.[23] The fungal gene family targeted in this study is the Glycoside Hydrolase 11 (GH11) family which encode *endo*-β-1,4-xylanases (E.C. 3.2.1.8) (CAZY Carbohydrate Active Enzymes database, http://www.cazy.org).[24] As xylan is the second most abundant polysaccharide in nature and one of the major structural polysaccharide in the plant cell wall, such enzymes have an obvious importance for soil ecology and for the degradation of plant hemicelluloses. A recent study also suggested that fungi contributed to most xylanase activity in soils.[25] Furthermore, GH11 enzymes are also abundantly used in different industrial processes.[26] GH11 genes are present in the genomes of numerous fungi, mainly Ascomycota and Basidiomycota, and at the start of this study, >300 sequences were publicly available. Furthermore, in a random shotgun sequencing of forest soil eukaryotic polyA-mRNAs, it was shown that GH11 transcripts occurred at a low frequency ranging from 0 to 1 per 10^4 sequences obtained.[22,27]

2. Materials and methods

2.1. Soil RNA extraction and cDNA synthesis

Four different forest soils from France and Italy were used in this study (see Supplementary Table S1 for sites and soils characteristics). At each site, between 30 (BEW) and 60 (BRH) sieved (2 mm) soil cores were mixed together to constitute composite samples which were stored at −75°C prior to RNA extraction. RNA was extracted from 4 to 48 g of soil using protocols adapted to each soil. RNA from the Puéchabon (PUE) sample was extracted according to Luis *et al.*[28] RNA from the

Breuil Spruce (BRE) and Breuil Beech (BRH) samples was extracted according to Damon *et al.*[29] RNA from the Berchidda (BEW) sample was extracted using the PowerSoil® Total RNA Isolation Kit (Mo Bio Laboratories), according to the manufacturer's instructions. All RNA samples were treated with RNase-free DNase I to remove residual DNA contaminations and quantified by spectrophotometry (ND-1000 NanoDrop®, Thermo Scientific).

Eukaryotic cDNAs were synthesized from 2 μg of total soil RNA by using the Mint-2 cDNA synthesis and amplification kit according to the manufacturer's instructions (Evrogen). First-strand synthesis was initiated at the RNA 3′ poly-A end using a modified poly-dT primer (CDS-4M). The number of PCR cycles (between 22 and 30) necessary for optimal synthesis of the double-stranded cDNA (dscDNA) was evaluated for each cDNA sample. As a result of using the Mint-2 kit, all amplified cDNAs were bordered at their 5′ end by the M1 sequence (AAGCAGTGGTATCAACGCAG AGT) and the *Sfi*IA restriction site (GGCCATTACGGCC) while, at their 3′ end, they were bordered by the *Sfi*IB restriction site (GGCCGAGGCGGCC) and the M1 sequence. dscDNA was purified by phenol–chloroform extraction, precipitated by 2.5 volume of ethanol and 0.1 volume of sodium acetate, resuspended in ultrapure water and quantified.

2.2. *Capture probe design and synthesis*

As in July 2012, all publicly available GH11 DNA coding sequences of eukaryotic origin were identified by BLAST searches[30] and collected from GenBank (http://www.ncbi.nlm.nih.gov/genbank/), the Joint Genome Institute database (http://jgi.doe.gov/), the Broad Institute genome database (http://www.broadinstitute.org/) and CAZy (http://www.cazy.org/). A set of thirty-five 31-mers, degenerate capture probes, targeting the catalytic domain of the encoded proteins (pfam no. PF00457, ~540 nucleotide long; Supplementary Fig. S1), was designed from a collection of 342 coding DNA sequences using the KASpOD software.[11] Individual probe coverage ranged from 7 to 54% of the 342 sequences, leading to a probe set coverage of 90% (four allowed mismatches).

The 35 oligonucleotide probes included the specific sequences $(X)_{31}$ targeting cDNAs encoding GH11 and adaptor sequences at each extremities for PCR amplification: ATCGCACCAGCGTGT-$(X)_{31}$-CACTGCGGCTCCT CA (Supplementary Table S2 and Fig. S1). Biotinylated RNA capture probes were prepared according to the two-step procedure of Gnirke *et al.*[31] In the first step, each single-stranded DNA probe was amplified by PCR using primers complementary to the 5′ and 3′ adaptors to allow double-strand DNA formation. In the second step, agarose gel-purified double-stranded DNA probes were converted into biotinylated RNA probes by *in vitro* transcription using the MEGAScript®T7 kit (Ambion) and biotin-dUTP (TeBu Bio). RNA probes were then mixed together in equimolar amounts.

2.3. *cDNA capture*

cDNA capture was carried out as described by Denonfoux *et al.*[9] and summarized in Supplementary Fig. S2. Briefly, 500 ng of heat denatured PCR-amplified cDNAs were hybridized to the equimolar mix of biotinylated RNA probes (500 ng) for 24 h at 65°C. Probe/cDNA hybrids were trapped by streptavidin-coated paramagnetic beads (Dynabeads® M-280 Streptavidin, Invitrogen). After different washing steps to remove unbound cDNAs, the captured cDNAs were eluted from the beads using 50 μl of 0.1 M NaOH at room temperature, neutralized with 70 μl of 1 M Tris–HCl, pH 7.5, and purified using the Qiaquick PCR purification kit (Qiagen).

Captured cDNAs were PCR amplified using primer M1 that binds at both 5′ and 3′ ends of the cDNAs. PCRs were set up using 5 μl of eluate, 200 μM of deoxynucleotides (dNTPs), 400 nM primer M1, 5 μl of reaction buffer 10× (Evrogen) and 1 μl of 50× Encyclo DNA polymerase (Evrogen) in a final volume of 50 μl. After an initial denaturation at 95°C for 1 min, cDNAs were amplified for 25 cycles comprising 15 s at 95°C, 20 s at 66°C and 3 min at 72°C. Ten independent amplifications were conducted for each sample. PCR products of the same sample were purified on QIAquick columns (Qiagen) and pooled. A second round of hybridization and PCR amplification was performed using each of the amplified cDNA samples obtained after the first hybridization capture. Purified products originating from the same cDNA sample were pooled together and quantified by spectrophotometry (NanoDrop™ 2000, Thermo Scientific). The DNA quality and size distribution of captured cDNA were assessed on an Agilent 2100 Bioanalyzer DNA 12000 chip (Agilent Technologies).

2.4. *Semi-quantitative PCR*

Enrichment in GH11 sequences at each step of the capture protocol was evaluated by semi-quantitative PCR using different quantities of cDNAs and GH11-fungal-specific degenerate primers GH11-F (GGVAAGG GITGGAAYCNNGG) and GH11-R (TGKCGRACIGACCA RTAYTG) amplifying a ±281-bp fragment (Luis P. *et al.*, unpublished). PCRs were performed using 10, 1, 0.1 or 0.01 ng cDNAs obtained before and after one or two cycles of hybridization capture. Twenty-five microlitres of PCR mixes contained 1 μl of template cDNA, 2.5 μl of 10× PCR buffer without Mg (Invitrogen), 1.5 mM $MgCl_2$, 0.8 mM of each dNTP, 0.5 μM of each primer and 1 U of *Taq* DNA polymerase

(Invitrogen). After an initial denaturation at 94°C for 3 min, GH11 gene fragments were amplified for 45 cycles comprising 45 s at 94°C, 45 s at 50°C and 2 min at 72°C. After a final elongation at 72°C for 10 min, 10 μl of PCR products were run in a 1.5% ethidium bromide-stained agarose gel.

2.5. *High-throughput sequencing*

Diversity of GH11 sequences at each step of the capture protocol was evaluated by high-throughput sequencing of GH11 PCR products obtained, as described above, using primers GH11-F and GH11-R. PCRs were performed using cDNAs obtained before and after one or two cycles of hybridization capture. Twenty-five microlitres of PCR mixes contained 10 ng of template cDNA, 2.5 μl of 10× PCR buffer without Mg (Invitrogen), 1.5 mM $MgCl_2$, 0.8 mM of each dNTP, 0.5 μM of each primer and 1.25 U of DNA polymerase (a 24:1 mix of Invitrogen *Taq* DNA polymerase and Biorad iProof polymerase). PCR cycling conditions were as described above. Five different PCRs were prepared and run in parallel for each cDNA sample. PCR products were first checked on 1.5% agarose gel before pooling together the five replicates and purification using the QIAquick PCR purification kit (Qiagen). Paired-end sequencing (2 × 250 bp) was carried out on an Illumina MiSeq sequencer (Fasteris, Switzerland).

Paired-end reads were assembled using PandaSeq v.2.5,[32] and all sequences containing unidentified nucleotide positions ('N') were filtered out. Primers and barcodes were removed using MOTHUR v.1.30.2.[33] UCHIME[34] was used for chimera detection, and sequence clusters were constructed at a 95% nucleotide sequence identity threshold. The most abundant representative sequence of each of the most abundant clusters, altogether encompassing >90% of the sequences, was translated into amino acid sequence using the ORF Finder tool of the Sequence Manipulation Suite[35] (http://www.bioinformatics.org/sms2/). Shannon diversity indices (H′) were calculated after rarefying the different data sets from the same soil to the same sequencing depth (i.e. the lowest sequencing depth of the three samples of each soil, Table 2). Venn diagrams were drawn using the BioVenn tool (http://www.cmbi.ru.nl/cdd/biovenn/).

2.6. *Full-length cDNA cloning and sequencing*

Amplified cDNAs obtained after two rounds of hybridization capture were digested by *Sfi*I (Fermentas), which recognizes two distinct *Sfi*IA and *Sfi*IB sites located at the 5′ and 3′ ends of the cDNAs, respectively. Digested cDNAs were then ligated to the *Sfi*I-digested pDR196-SfiI-Kan yeast expression vector[36] modified to contain two *Sfi*IA and *Sfi*IB sites, downstream of the *Saccharomyces cerevisiae PMA1* promoter, thus allowing the directional cloning and potential constitutive expression of the cDNAs in yeast.

Several transformed, kanamycin-resistant *Escherichia coli* (One Shot® TOP10 strain, Invitrogen) colonies from each sample were first randomly selected and subjected to colony PCR using the GH11-F and GH11-R primers to detect the presence of a GH11 cDNA insert. cDNA inserts from PCR-positive bacterial colonies were entirely sequenced by BIOFIDAL (Villeurbanne, France) using a PMA1 primer (CTCTCTTTTATACACACATTC) and additional internal primers when necessary.

2.7. *Plasmid library construction, yeast transformation and functional screening*

For each cDNA sample, a minimum of 2,000 independent kanamycin-resistant transformed *E. coli* colonies were pooled together for plasmid extraction using the alkaline lysis method.[37] Aliquot samples of each plasmid library were used to transform the *S. cerevisiae* strain DSY-5 (*MATα leu2 trp1 ura3-52 his3::PGAL1-GAL4 pep4 prb1-1122*; Dualsystems Biotech) using a standard lithium acetate protocol.[38] Transformed yeasts were selected on a solid yeast nitrogen base (YNB) minimal medium supplemented with glucose (2%) and amino acids, but lacking uracil. YNB agar plates were overlaid by a thin layer of the same medium containing 4 mg l^{-1} of AZCL-xylan (Megazyme), a substrate specific for *endo*-xylanases. Plates were incubated at 30°C. Yeast colonies producing a secreted *endo*-xylanase were surrounded by a dark blue halo resulting from the hydrolysis of AZCL-xylan.

For each sample, several yeast colonies positive for *endo*-xylanase activity were picked, lysed at 95°C for 10 min in 3 μl of 20 mM NaOH and the pDR196 insert amplified by PCR using primers PMA1 and ADH (GCGAATTTCTTATGATTTATG). PCR products were sequenced by BIOFIDAL using the PMA1 primer.

2.8. *Phylogenetic analyses*

Sequences obtained from plasmid inserts were manually edited and corrected. Deduced amino acid sequences were aligned using MUSCLE[39] to GH11 amino acid sequences obtained from public databases. Maximum likelihood phylogeny analyses were generated with the PhyML 3.0 program using the WAG substitution model as implemented in SeaView v. 4.[40] Phylogenetic trees were drawn in MEGA v. 6.[41]

2.9. *Sequence accessibility*

Sequences from plasmid inserts are available in the EBI/DDJB/GenBank databases under accession Nos. LK932029-LK932091. Illumina MiSeq sequence reads have been deposited in the Sequence Read Archive of the EBI database under study no. PRJEB6672.

3. Results

3.1. *GH11 cDNA capture*

As in July 2012, we identified and collected 342 full-length eukaryotic GH11 DNA coding sequences from public databases, from 113 fungal species and from 2 non-fungal ones. Seventy-two percent of these sequences were from Ascomycotina (85 species), 20% from Basidiomycotina (26 species) and 7% from other taxonomic groups. Prevalence of sequences from Ascomycotina is likely to reflect a greater genome sequencing effort in this taxonomic group, rather than a higher occurrence of the GH11 family among Ascomycotina.[23] Among the publicly available sequences, those putatively full-length sequences ranged in size from 639 to 2,099 bp. Occurrence of carbohydrate-binding motives or of C-terminal, non-catalytic extensions in the encoded polypeptides accounted for most of these size variations. The 35 degenerate capture probes were exclusively designed on the shared ca. 540-bp-long conserved catalytic domain and were susceptible to hybridize to 90% of the collected sequences.

SHS was performed on cDNAs synthesized from polyadenylated mRNAs extracted from four different forest soils. Electrophoregrams of all cDNAs recovered after two successive rounds of capture were characterized by a background smear of which emerged discrete bands ranging in size from 300 to 1,500 bp (Fig. 1).

Successful enrichment in GH11 sequences along the capture protocol was demonstrated by semi-quantitative PCR using GH11-specific PCR primers and different quantities of cDNA in the PCRs (from 10 to 0.01 ng). As illustrated in Fig. 2 for the Breuil beech forest (BRH sample) and for the other soil samples discussed in Supplementary Fig. S3, clear positive amplification of a GH11 fragment after two rounds of capture was always obtained using the lowest quantity of cDNA (0.01 ng), whereas no amplification could be observed for the same amount of cDNA prior to SHS.

3.2. *Cloning, sequencing and heterologous expression of captured cDNA*

Captured cDNAs in the range of 700–1,500 bp were cloned into the pDR196 *E. coli/S. cerevisiae* shuttle expression vector to constitute four soil-specific GH11-enriched plasmid libraries (Table 1). Forty recombinant colonies per library were randomly screened by PCR using GH11-specific primers to evaluate the percentage of GH11-containing recombinant plasmids. Efficient enrichment occurred for all libraries with 80 to >90% of positive clones (Table 1). Among the 55 fully sequenced plasmid inserts from PCR- positive

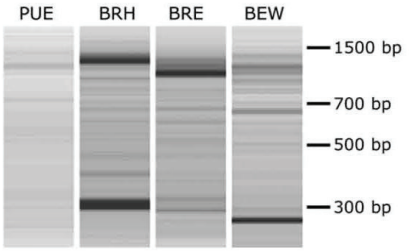

Figure 1. Electrophoretic separation of cDNAs obtained following two consecutive solution hybridization selection. Captured cDNAs from the four soil samples PUE, BRH, BRE and BEW were run on an Agilent DNA 12000 microfluidic chip. Each band could encompass one or several unique but abundant GH11 cDNAs.

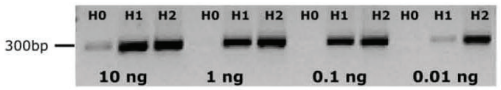

Figure 2. Semi-quantitative PCR amplification of a 281-bp GH11 fragment using different quantities (from 10 to 0.01 ng) of BRH cDNA obtained before (H0) and after one (H1) or two (H2) cycles of hybridization. Before capture, PCR products could only be obtained using 10 ng of input cDNA. Amplifications of the PUE, BRE and BEW samples are illustrated in Supplementary Fig. S3.

colonies, all but two indeed corresponded to GH11 sequences (Table 1). Seventy-two percent of the sequences encoded putatively full-length GH11 polypeptides based on alignment length to known GH11 polypeptides and the presence of in-frame putative start and stop codons. Out of them, 15% were characterized by the presence of a family 1 carbohydrate-binding domain (CBM1) in a C-terminal position.

Functional screening using *S. cerevisiae* was conducted on the four GH11-enriched plasmid libraries by plating the recombinant yeasts onto a medium supplemented with an *endo*-xylanase-specific colour reagent (AZCL-xylan). Depending on the library, between 1.5 (sample PUE) and 25% (sample BRH) of the transformed yeast colonies developed a dark blue halo demonstrating secretion of a functional *endo*-xylanase (Supplementary Fig. S4). All 11 sequenced plasmid inserts from these xylanase-positive yeast colonies encoded GH11 proteins (ranged between 221 and 289 amino acids in length); 5 of them had already been identified among sequences obtained from bacterial colonies and 4 had a C-terminal CBM1 domain. The percent sequence identity between the catalytic domain of the selected functional proteins and the catalytic domain of their closest Blastp hits in GenBank ranged between 69% (81% similarity) and 87% (94% similarity).

Table 1. Cloning and characterization of captured GH11 cDNAs

Samples	PUE	BRH	BRE	BEW
No. of captured cDNAs cloned in *Escherichia coli*	6,770	2,020	5,720	5,880
No. of *E. coli* colonies screened by PCR	40	40	40	40
Positive amplification of a GH11 fragment (%)	37 (92.5)	33 (82.5)	35 (87.5)	36 (90)
No. of inserts sequenced	12	13	16	14
No. of GH11 inserts (%)	11 (92)	12 (92)	16 (100)	14 (100)
No. of putative full-length GH11 (%)	9 (82)	9 (75)	11 (69)	9 (64)
% of *endo*-xylanase-positive yeast colonies	1.5	25	12	6

Table 2. Summary statistics from Illumina MiSeq sequencing of GH11 PCR fragments amplified, for each four cDNA samples, before (H0) or after one (H1) or two (H2) hybridization capture

Sample	Total no. of sequences	Total no. of clusters[a] (95%)	No. of clusters encompassing ≥90% of the sequences	Shannon diversity index (H')[b]	No. of shared clusters between H0−H1−H2[b]
PUE_H0	12,960	298	52 (17%)	3.819	70 (11%)
PUE_H1	24,565	227	51 (22%)	4.015	
PUE_H2	25,053	291	46 (16%)	3.912	
BRE_H0	13,538	87	9 (10%)	2.254	11 (5%)
BRE_H1	42,000	140	5 (4%)	1.651	
BRE_H2	46,626	112	6 (5%)	1.73	
BRH_H0	2,765	26	3 (12%)	1.061	5 (4%)
BRH_H1	28,366	51	3 (6%)	1.234	
BRH_H2	17,322	159	18 (11%)	2.135	
BEW_H0	41,799	214	15 (7%)	2.761	38 (6%)
BEW_H1	42,308	249	10 (4%)	2.496	
BEW_H2	36,859	205	6 (3%)	2.196	

[a]Including singletons.
[b]Shannon diversity indices and shared clusters were calculated after rarefying the different data sets from the same soil to the same sequencing depth of 12,960, 13,538, 2,765 and 36,859 sequences for PUE, BRE, BRH and BEW, respectively.

3.3. *Selectivity of the SHS GH11 capture*

To evaluate the diversity of GH11 sequences at each step of the capture protocol, we performed a high-throughput Illumina MiSeq sequencing of GH11 amplicons obtained from all four cDNA samples, prior (H0) and after one (H1) or two (H2) cycles of SHS capture. Paired-end sequence reads were assembled to reconstitute the ca. 281-bp-long amplicons. Altogether, the total data set contained 334,161 full-length amplicon sequences that were clustered at a 95% nucleotide sequence identity threshold to produce a total number of 1,458 clusters, of which 1,001 (69%) were singletons (data summarized in Table 2 for each sample). Each of the 12 sequence data sets (4 cDNA samples × the 3 steps of the SHS) was characterized by few dominant clusters encompassing most of the sequences and a large number of clusters each containing a few, or even a single, sequences (illustrated in Fig. 3A for the PUE sample). None of the sequences obtained were identical to sequences deposited in databases. Only 17 of the sequence clusters, of which 14 exclusively from the BEW site, were >90% identical (maximum value of 97.5%) at the nucleotide level over their entire length to GH11 genes from either the Basidiomycota *Tulasnella calospora* or the Ascomycota *Nectria haematococca* and *Pyrenophora teres*.

Figure 3 also showed that the most abundant sequence clusters obtained after one (H1) and two (H2) cycles of capture did not, for a majority of them, correspond to the most abundant clusters present before capture (H0). Venn diagrams drawn using only these most prominent sequence clusters, encompassing altogether 90−93% of sample sequences, showed that there existed a larger overlap between the post-capture samples H1 and H2 than between the pre-capture samples H0 and H1 or H2 (Fig. 3B). This trend was observed, to some extent, for samples BEW, BRE and PUE, but not for the BRH one which differed from the others by the dominance of only three clusters in the H0 cDNA pool which encompassed 90% of the

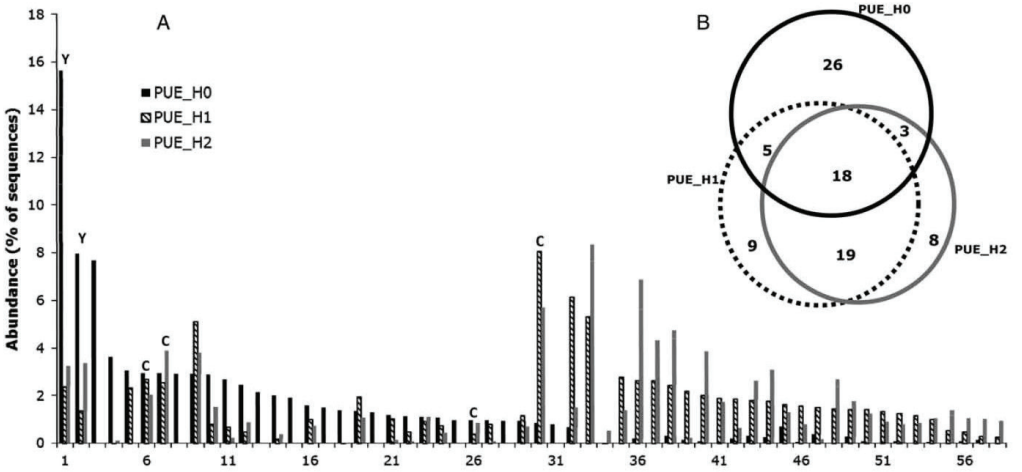

Figure 3. Selectivity of the SHS capture. (A) Rank-abundance distribution of the most abundant GH11 nucleotide sequence clusters identified before (H0), or after one (H1) or two (H2) cycles of hybridization on the PUE cDNAs. Only clusters encompassing 80% of the sequences in the H0, H1 or H2 samples are shown. 'C' or 'Y' letters above bars indicate sequences obtained by random sequencing of plasmid inserts or which could be functionally expressed in yeast, respectively. (B) Venn diagram showing the number of unique or shared GH11 sequence clusters, before (H0), or after one (H1) or two (H2) cycles of hybridization on the PUE cDNAs. As in (A), only the most abundant clusters, encompassing 90% of the sequences, were used for the calculation. GH11 PCR sequences were clustered using a nucleotide sequence identity threshold of 95%. Similar Venn diagrams for the BRH, BRE and BEW samples are illustrated in Supplementary Fig. S5.

sequenced reads (Supplementary Fig. S5). Despite these apparent differences in sequence distribution between the pre-capture H0 and the post-capture H1 and H2 samples, sequence diversity indices, such as the Shannon index, did not differ between the pre- and post-capture sequence pools (Table 2, with the exception of the BRH sample). Between 2.7% (BRE and BEW) and 15% (PUE and BRH) of the sequence clusters were shared between two sites. Eight sequence clusters were identified in all four studied sites.

To address the phylogenetic diversity of the captured sequences, we first produced an amino acid sequence alignment of 62 known GH11 proteins representative of the phylogenetic diversity of this gene family. To this alignment, we added the GH11 sequences obtained by the random sequencing of plasmid inserts, the sequences producing a functional enzyme in yeast and the sequences representative of the most abundant Illumina sequence clusters before (H0) or after (H1 and H2) SHS capture. The GH11 family is a highly diversified and fast-evolving gene family and phylogenies based either on full-length protein sequence alignments or on partial alignments, as in the present case, clearly do not reflect the species phylogenies and comprise very few well-supported internal branches (Fig. 4). Phylogenetic trees obtained for sequences from the four studied soils (Fig. 4; Supplementary Fig. S6) all clearly showed that the captured sequences were distributed over the entire reference tree.

4. Discussion

The results obtained clearly demonstrate that SHS represents a powerful strategy to select full-length cDNAs, representative of a specific gene family, originally diluted in a highly complex metatranscriptomic sequence pool. This protocol was successfully implemented on four different forest soil RNA samples. Based on previous estimates of the frequency of GH11 sequences among eukaryotic cDNA for two of the soils used in this study (BRE and BRH),[22] two successive cycles of SHS have the potential to enrich specific cDNA sequences by a factor of at least 10^4. As suggested by the results of the semi-quantitative PCR, in some cases (e.g. the PUE sample, Supplementary Fig. S3), one cycle of capture may be sufficient to get a maximum level of enrichment, while in other cases two cycles seem required (e.g. the BRH sample, Fig. 2).

Sequence analysis of PCR fragments amplified from pre- or post-capture cDNAs demonstrated that capture succeeded in selecting both a large number and phylogenetically diverse representatives of the selected gene family. Furthermore, none of the captured sequences appeared to be identical to already known ones which we originally used for probe design. Capture could however preferentially select sequences that were not necessarily among the most abundant in the original cDNA pool. This should be evaluated in the future by quantitative PCR. Despite

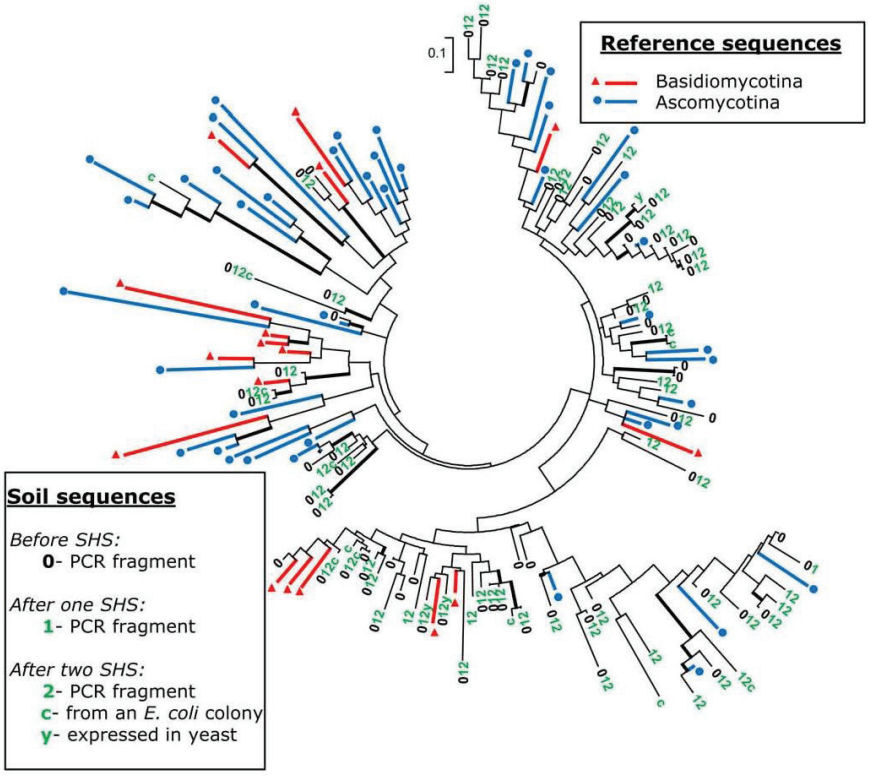

Figure 4. Phylogenetic diversity of the GH11 partial amino acid sequences obtained from PUE cDNA samples. 0, 1 and 2 translated PCR sequences obtained before or after one or two cycles of hybridization. PUE sequences are scattered over the entire tree that includes representative reference sequences from Ascomycota and Basidiomycota. c, sequences obtained from *Escherichia coli* clones; y, sequences functionally expressed in yeast clones. PhyML tree calculation was based on an alignment of ca. 80-amino-acid-long GH11 partial sequences. Thicker internal black branches indicate bootstrap value ≥60% (1,000 replications). Full species names and accession numbers of the reference sequences are given in Supplementary Fig. S6A. Similar trees drawn using the sequences from sites BRE, BRH and BEW are illustrated in Supplementary Fig. S6 B, C and D, respectively.

explorative probe design strategy, publicly available homologous sequences at the start of the study greatly influence the capture selectivity. Probe sets utilized to capture a given biomarker should therefore be upgraded regularly, taking into account newly deposited sequences.

Thanks to the ever increasing number of published fungal genomes, representative of the phylogenetic diversity of this taxonomic group; explorative probe design strategies could be carried out to unravel the metabolic capacities of these microorganisms within different ecosystems. Besides GH11 sequences, SHS capture can be implemented for any other gene family of interest, allowing a comprehensive taxonomic or functional description of the studied microbial community. As mentioned in the introduction, sequence capture presents the advantage over PCR to give access to the full-length gene sequence, including facultative modules, not always associated to the studied catalytic domain. This was indeed the case for the GH11, for which we estimated that 72% of the captured sequences were full length and that 15% of them processed a C-terminal, fungal-specific, CBM1 module (see the CAZy database, http://www.cazy.org). A discrepancy however existed between the estimated fraction of full-length captured GH11 cDNA and the systematically lower fraction of cDNAs which produced a functional enzyme upon expression in *S. cerevisiae*. The absence of expression in yeast can be attributed to a number of independent factors such as bias in codon usage, non-recognition by *S. cerevisiae* of the protein signal peptide necessary for correct secretion, protein misfolding or hyperglycosylation. Some of these problems could be addressed by using expression

plasmids including a yeast signal peptide downstream of the cloning site and/or by using a different yeast species for protein production.

Sequencing of PCR fragments amplified from captured cDNAs also indicate that altogether the four captured cDNA samples obtained in this single study encompass a greater number of novel and different GH11 sequences than have been deposited and are available in public databases over several decades. This observation should promote the use of cDNA sequence capture (i) as a complementary approach to PCR to explore and quantify the extent of eukaryotic functional diversity in complex environments, but also (ii) as a powerful tool in environmental biotechnology to efficiently screen for enzyme variants with novel biochemical properties.

Acknowledgements: We would like to thank Richard Joffre, Jacques Ranger and Alberto Orgiazzi for soil sampling at the Puéchabon, Breuil and Berchidda sites, respectively. Audrey Dubost and Stefano Ghignone contributed to bioinformatic analyses and Jérémie Denonfoux to gene capture. We acknowledge the JGI of the US Department of Energy and the Broad Institute for making available genome data prior to their publication.

Supplementary Data: Supplementary Data are available at www.dnaresearch.oxfordjournals.org.

Funding

C.B. was supported by the University of Turin and the région Rhône-Alpes (CMIRA program); C.R. received a graduate grant from the Ministère de l'Enseignement Supérieur et de la Recherche; N.P. was funded by the Direction Générale de l'Armement and E.P. by the Agence Nationale pour la Recherche. Work was financed by the CNRS-INSU ECCO Microbien program, the INRA métaprogramme M2E (project Metascreen), project ANR 09-GENM-033-001 (Eumetasol); EU-project 'EcoFINDERS' No. 264465 and local funding by the University of Turin (ex-60%). Funding to pay the Open Access publication charges for this article was provided by the Centre national de la recherche scientifique (CNRS).

References

1. Tringe, S.G. and Rubin, E.M. 2005, Metagenomics: DNA sequencing of environmental samples, *Nat. Rev. Genet.*, **6**, 805–14.
2. Simon, C. and Daniel, R. 2011, Metagenomic analyses: past and future trends, *Appl. Environ. Microbiol.*, **77**, 1153–61.
3. Howe, A.C., Jansson, J.K., Malfatti, S.A., Tringe, S.G., Tiedje, J.M. and Brown, C.T. 2014, Tackling soil diversity with the assembly of large, complex metagenomes, *Proc. Natl. Acad. Sci. USA*, **111**, 4904–9.
4. Voříšková, J. and Baldrian, P. 2013, Fungal community on decomposing leaf litter undergoes rapid successional changes, *ISME J.*, **7**, 477–86.
5. Luis, P., Kellner, H., Zimdars, B., Langer, U., Martin, F. and Buscot, F. 2005, Patchiness and spatial distribution of laccase genes of ectomycorrhizal, saprotrophic, and unknown basidiomycetes in the upper horizons of a mixed forest cambisol, *Microb. Ecol.*, **50**, 570–9.
6. Kellner, H., Luis, P., Pecyna, M.J., et al. 2014, Widespread occurrence of expressed fungal secretory peroxidases in forest soils, *PLoS ONE*, **9**, e95557.
7. Kellner, H., Zak, D.R. and Vandenbol, M. 2010, Fungi unearthed: transcripts encoding lignocellulolytic and chitinolytic enzymes in forest soil, *PLoS ONE*, **5**, e10971.
8. Hong, S., Bunge, J., Leslin, C., Jeon, S. and Epstein, S.S. 2009, Polymerase chain reaction primers miss half of rRNA microbial diversity, *ISME J.*, **3**, 1365–73.
9. Denonfoux, J., Parisot, N., Dugat-Bony, E., et al. 2013, Gene capture coupled to high-throughput sequencing as a strategy for targeted metagenome exploration DNA, *DNA Res.*, **20**, 185–96.
10. Dugat-Bony, E., Missaoui, M., Peyretaillade, E., et al. 2011, HiSpOD: probe design for functional DNA microarrays, *Bioinformatics*, **27**, 641–8.
11. Parisot, N., Denonfoux, J., Dugat-Bony, E., Peyret, P. and Peyretaillade, E. 2012, KASpOD-a web service for highly specific and explorative oligonucleotide design, *Bioinformatics*, **28**, 3161–2.
12. Bailly, J., Fraissinet-Tachet, L., Verner, M.C., et al. 2007, Soil eukaryotic functional diversity, a metatranscriptomic approach, *ISME J.*, **1**, 632–42.
13. Kellner, H., Luis, P., Portetelle, D. and Vandenbol, M. 2011, Screening of a soil metatranscriptomic library by functional complementation of *Saccharomyces cerevisiae* mutants, *Microbiol. Res.*, **166**, 360–8.
14. Damon, C., Vallon, L., Zimmermann, S., et al. 2011, A novel fungal family of oligopeptide transporters identified by functional metatranscriptomics of soil eukaryotes, *ISME J.*, **5**, 1871–80.
15. Lehembre, F., Doillon, D., David, E., et al. 2013, Soil metatranscriptomics for mining eukaryotic heavy metal resistance genes, *Environ. Microbiol.*, **15**, 2829–40.
16. Takasaki, K., Miura, T., Kanno, M., et al. 2013, Discovery of glycoside hydrolase enzymes in an avicel-adapted forest soil fungal community by a metatranscriptomic approach, *PLoS ONE*, **8**, e55485.
17. Bates, S.T., Clemente, J.C., Flores, G.E., et al. 2013, Global biogeography of highly diverse protistan communities in soil, *ISME J.*, **7**, 652–9.
18. Taylor, D.L., Hollingsworth, T.N., McFarland, J.W., Lennon, N.J., Nusbaum, C. and Ruess, R.W. 2014, A first comprehensive census of fungi in soil reveals both hyperdiversity and fine-scale niche partitioning, *Ecol. Monogr.*, **84**, 3–20.
19. Schneider, T., Keiblinger, K.M., Schmid, E., et al. 2012, Who is who in litter decomposition? Metaproteomics reveals

major microbial players and their biogeochemical functions, *ISME J.*, **6**, 1749–62.
20. Stursová, M., Zifčáková, L., Leigh, M.B., Burgess, R. and Baldrian, P. 2012, Cellulose utilization in forest litter and soil: identification of bacterial and fungal decomposers, *FEMS Microbiol. Ecol.*, **80**, 735–46.
21. Demain, A.L. and Dana, C.A. 2007, The business of biotechnology, *Ind. Biotechnol.*, **3**, 269–83.
22. Damon, C., Lehembre, F., Oger-Desfeux, C., et al. 2012, Metatranscriptomics reveals the diversity of genes expressed by eukaryotes in forest soils, *PloS ONE*, **7**, e28967.
23. Grigoriev, I.V., Nikitin, R., Haridas, S., et al. 2014, MycoCosm portal: gearing up for 1000 fungal genomes, *Nucleic Acids Res.*, **42**, D699–704.
24. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M. and Henrissat, B. 2014, The Carbohydrate-active enzymes database (CAZy) in 2013, *Nucleic Acids Res.*, **42**, D490–495.
25. Margesin, R., Minerbi, S. and Schinner, F. 2014, Long term monitoring of soil microbiological activities in two forest sites in South tyrol in the italian alps, *Microb. Environ.*, **29**, 277–85.
26. Paës, G., Berrin, J.G. and Beaugrand, J. 2012, GH11 xylanases: structure/function/properties relationships and applications, *Biotechnol. Adv.*, **30**, 564–92.
27. Kuramae, E.E., Hillekens, R.H., de Hollander, M., et al. 2013, Structural and functional variation in soil fungal communities associated with litter bags containing maize leaf, *FEMS Microbiol. Ecol.*, **84**, 519–31.
28. Luis, P., Kellner, H., Martin, F. and Buscot, F. 2005, A molecular method to evaluate Basidiomycete laccase gene expression in forest soils, *Geoderma*, **128**, 18–27.
29. Damon, C., Barroso, G., Férandon, C., Ranger, J., Fraissinet-Tachet, L. and Marmeisse, R. 2010, Performance of the COX1 gene as a marker for the study of metabolically active Pezizomycotina and Agaricomycetes fungal communities from the analysis of soil RNA, *FEMS Microbiol. Ecol.*, **74**, 693–705.
30. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990, Basic local alignment search tool, *J. Mol. Biol.*, **215**, 403–10.
31. Gnirke, A., Melnikov, A., Maguire, J., et al. 2009, Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing, *Nat. Biotechnol.*, **27**, 182–9.
32. Masella, A.P., Bartram, A.K., Truszkowski, J.M., Brown, D.G. and Neufeld, J.D. 2012, PANDAseq: paired-end assembler for illumina sequences, *BMC Bioinformatics*, **13**, 31.
33. Schloss, P.D., Westcott, S.L., Ryabin, T., et al. 2009, Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities, *Appl. Environ. Microbiol.*, **75**, 7537–41.
34. Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C. and Knight, R. 2011, UCHIME improves sensitivity and speed of chimera detection, *Bioinformatics*, **27**, 2194–200.
35. Stothard, P. 2000, The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences, *Biotechniques*, **28**, 1102–4.
36. Rentsch, D., Laloi, M., Rouhara, I., Schmelzer, E., Delrot, S. and Frommer, W.B. 1995, NTR1 encodes a high affinity oligopeptide transporter in Arabidopsis, *FEBS Lett.*, **370**, 264–8.
37. Sambrook, J. and Russell, D.W. 2001, *Molecular cloning: a laboratory manual*. 3rd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
38. Rose, M., Winston, F. and Hieter, P. 1990, *Methods in yeast genetics: a laboratory course manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
39. Edgar, R.C. 2004, MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acids Res.*, **32**, 1792–7.
40. Gouy, M., Guindon, S. and Gascuel, O. 2010, SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building, *Mol. Biol. Evol.*, **27**, 221–4.
41. Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. 2013, MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0, *Mol. Biol. Evol.*, **30**, 2725–9.

Supplementary Data

Sample name	*Geographic position*	*Mean elevation*	*Climate*	*Bedrock*	*Soil type*[a]	*pH*	*Organic matter (g/kg)*	*Predominant tree species*
Puéchabon (PUE)	Southern France 43° 44′ 29″ N, 3° 35′ 45″ E	270 m	Mediterranean	Jurassic limestone	Silty clay loam texture[1]	7.3	205	*Quercus ilex*
Breuil (BRH)	Central France 47° 18′ 10″ N, 4° 4′ 44″ E	640 m	Continental	Granitic rock	Alocrisol[2]	3.8	91.1	*Fagus sylvatica*
Breuil (BRE)						3.7	100	*Picea abies*
Berchidda (BEW)	Island of Sardinia (Italy) 40° 30′ 13.37″ N, 9° 47′ 00.56″ E	285 m	Mediterranean	Paleozoic intrusive rock (monzogranites)	Typic Dystroxerepts, with sandy loam texture[3]	6.2	33-59	*Quercus suber*

Table S1. Studied soils

[a]USDA classification, [1]Rambal *et al.*2003, [2]Ranger *et al.* 2004, [3]Lagomarsino *et al.* 2011

[1]Rambal, S., Ourcival, J.-M., Joffre, R., Mouillot, F., Nouvellon, Y., Reichstein, M. and Rocheteau, A. 2003, Drought controls over conductance and assimilation of a Mediterranean evergreen ecosystem: scaling from leaf to canopy. *Global Change Biology*, 9,1813–1824.

[2]Ranger, J., Andreux, J., Bienaimé, S.F, *et al.* 2004. Effet des substitutions d'essence sur le fonctionnementorgano-minéral de l'écosystème forestier, sur les communautés microbiennes et sur la diversité des communautés fongiques mycorhiziennes et saprophytes (cas du dispositif expérimental de Breuil-Morvan). Final report of contract INRA-GIP Ecofor 2001-24.Champenoux, France: N° INRA 1502A, INRA BEF Nancy.

[3]Lagomarsino, A., Benedetti, A., Marinari, S., *et al.* 2011, Soil organic C variability and microbial functions in a Mediterranean agro-forest ecosystem. *Biology and Fertility of Soils*, 47(3), 283–291.

Table S2. List of the degenerate probes used in this study to target sequences encoding eukaryotic GH11 (DNA sequences with 5' and 3'adapters)

Name	*Degeneracy*	*Probe sequence (5'-3')*
GH11_1	256	atcgcaccagcgtgtTAYGGYTGGACCMGMAACCCKCTSATYGARTcactgcggctcctca
GH11_2	192	atcgcaccagcgtgtGAARCCCKCTBRTYGARTACTACATYGTCGAcactgcggctcctca
GH11_3	64	atcgcaccagcgtgtGGACSACCARCCCGCTGGTYGARTAYTACRTcactgcggctcctca
GH11_4	192	atcgcaccagcgtgtTGATHGARTACTACATCGTYGARRMCTWCGGcactgcggctcctca
GH11_5	216	atcgcaccagcgtgtGCGGYAACYTYGTCGGHGGHAAGGGHTGGAAcactgcggctcctca
GH11_6	144	atcgcaccagcgtgtGGHAACTKSGTCGGYGGDAARGGATGGAACCcactgcggctcctca
GH11_7	128	atcgcaccagcgtgtTCGAGTACTAYATCGTYGARWSCTACGGYWCcactgcggctcctca
GH11_8	192	atcgcaccagcgtgtTGGCRACYTTGTYGSYGGMAAGGGVTGGAACcactgcggctcctca
GH11_9	144	atcgcaccagcgtgtACGGBTWCTWCTAYTCCTTCTGGACBGAYGGcactgcggctcctca
GH11_10	128	atcgcaccagcgtgtCTSKCYRTCTACGGCTGGRCSACSAACCCGCcactgcggctcctca
GH11_11	192	atcgcaccagcgtgtCCGGCAACDSCTACMTSKCCRYCTACGGCTGcactgcggctcctca
GH11_12	48	atcgcaccagcgtgtGTACTACRTCGTYGAGTCVTWCGGCACSTACcactgcggctcctca
GH11_13	64	atcgcaccagcgtgtCGGMTGGACCMSCRGYCCCCTCRTCGAGTACcactgcggctcctca
GH11_14	192	atcgcaccagcgtgtTACRTBGTCGARTCCTWCGGTRMCTACRACCcactgcggctcctca
GH11_15	192	atcgcaccagcgtgtCCBCTKATCGAGTACTAYRTCRTYGAGARCTcactgcggctcctca
GH11_16	128	atcgcaccagcgtgtCGTCGARTACTACATCSTSGAAWMYTWCGGCcactgcggctcctca
GH11_17	256	atcgcaccagcgtgtCCCRSACYTTCMASCAGTACTGGKCYRTCCGcactgcggctcctca
GH11_18	256	atcgcaccagcgtgtACTGGRTYGGYGGWAARGGRTGGAAYCCYGGcactgcggctcctca
GH11_19	256	atcgcaccagcgtgtAACYMACRACGGYTACTWCTACTCSTKSTGGcactgcggctcctca
GH11_20	128	atcgcaccagcgtgtAACGGCAAYKCYTACCTCKSMRTCTACGGCTcactgcggctcctca
GH11_21	128	atcgcaccagcgtgtCATYGTCGAGAASTWYGGCRMCTACRACCCCcactgcggctcctca
GH11_22	192	atcgcaccagcgtgtCACYTTYYABCAGTACTKGTCYGTCCGYCAGcactgcggctcctca
GH11_23	192	atcgcaccagcgtgtTACTAYATCGTCGARTCVTWCGGYTCSTACRcactgcggctcctca
GH11_24	128	atcgcaccagcgtgtATYGTCGAGAASTWYGGCRMCTACRACCCCTcactgcggctcctca
GH11_25	128	atcgcaccagcgtgtACTACATYSTSGARTCCTACGGCWCMTAYAAcactgcggctcctca
GH11_26	192	atcgcaccagcgtgtGCACYTTYCAGCAGTACTGGKCYRTCCGCMVcactgcggctcctca
GH11_27	128	atcgcaccagcgtgtCCCWGRYYGARTACTACRTCGTCGAGTCSTAcactgcggctcctca
GH11_28	256	atcgcaccagcgtgtCAGTWCTGGTCYGTYCGYCAGRACMASCGCWcactgcggctcctca
GH11_29	64	atcgcaccagcgtgtCGRTRACTTCGTCGKYGGAAAGGGMTGGARGcactgcggctcctca
GH11_30	108	atcgcaccagcgtgtCCTCBATCSABGGYACHCAGACCTTCCAGCAcactgcggctcctca
GH11_31	128	atcgcaccagcgtgtAGCACYAACCCYCTTGTYGARTACTAYRTCRcactgcggctcctca
GH11_32	144	atcgcaccagcgtgtAGCCBTCCATCRTCGGMACBGCCACCTTYWAcactgcggctcctca
GH11_33	192	atcgcaccagcgtgtCTCSTWCCTBKCYGTSTACGGCTGGRTCAACcactgcggctcctca
GH11_34	64	atcgcaccagcgtgtCTGCRACAACTTYGTYGCYGGMAAGGGMTGGcactgcggctcctca
GH11_35	192	atcgcaccagcgtgtTCGTYGTYGGYGTTGGCTGGRVMMCTGGATCcactgcggctcctca

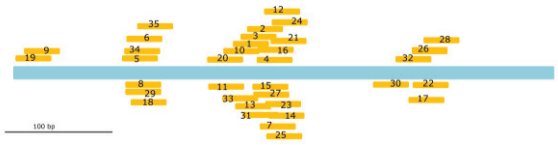

Figure S1. Position of the 35 capture probes along the GH11 catalytic domain.

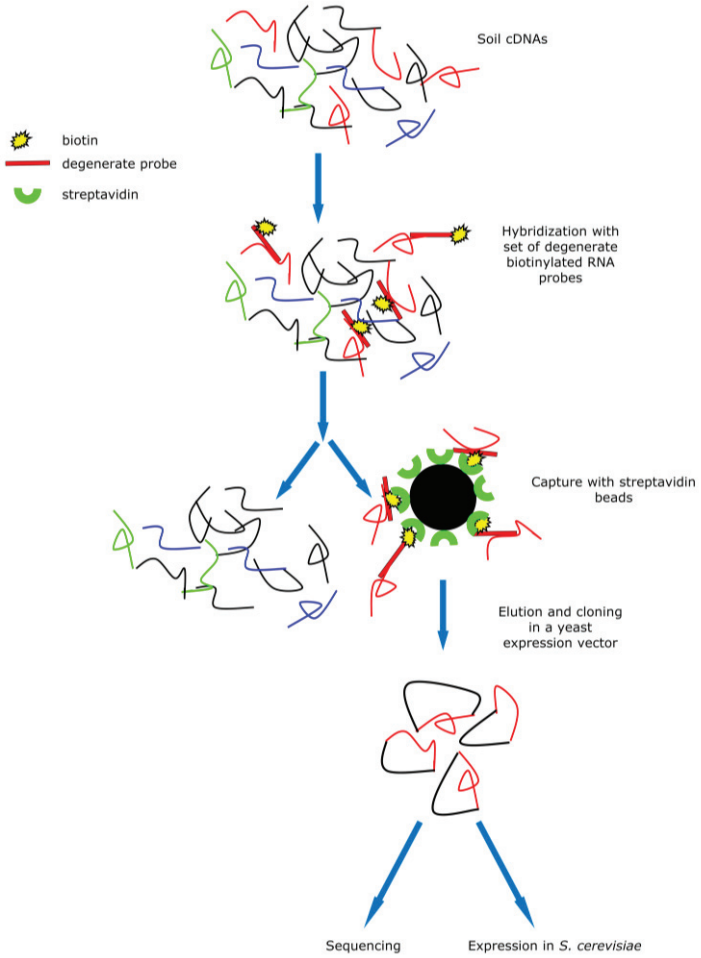

Figure S2. Overview of the SHS capture method implemented in the present study. The first two steps were performed twice consecutively.

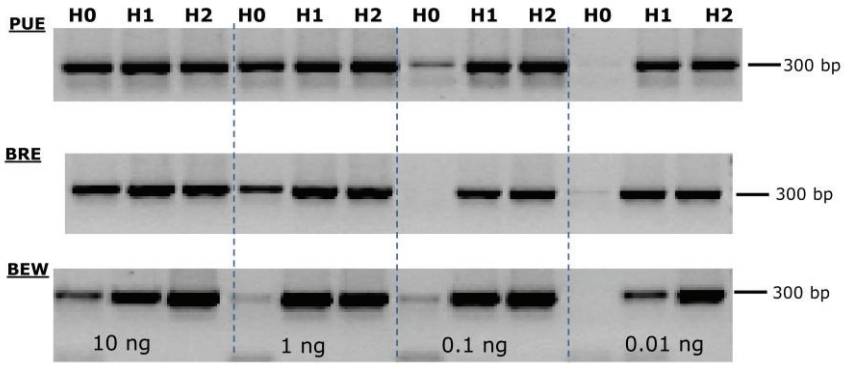

Figure S3. Semi-quantitative PCR amplification of a 281 bp GH11 fragment using different quantities (from 10 to 0.01 ng) of PUE, BRE or BEW cDNAs obtained before (H0) and after one (H1) or two (H2) cycles of hybridization.

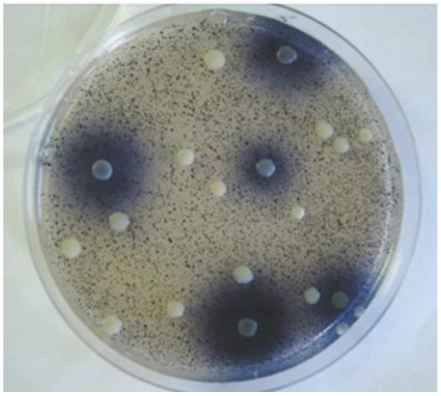

Figure S4. Several of the yeast colonies transformed with the plasmid library prepared from captured BRH cDNAs express a functional secreted endo-xylanase. Following transformation, DSY-5 yeast cells were plated on a selective medium without uracil and containing AZCL-xylan, an endoxylanase-specific substrate, whose degradation leads to the release of a dark blue dye.

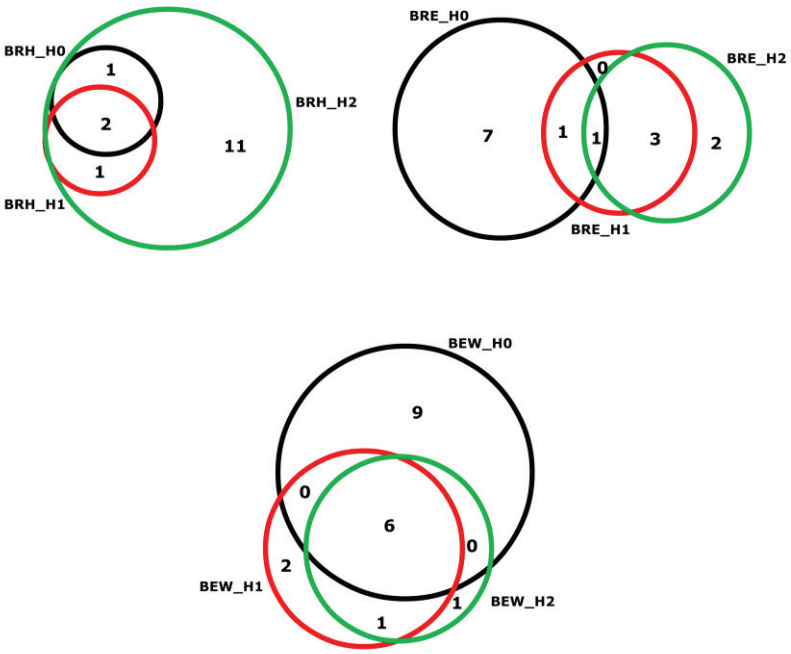

Figure S5. Selectivity of the Solution Hybrid Selection (SHS) capture. Venn diagram showing the number of unique or shared GH11 sequence clusters, before (H0), or after one (H1) or two (H2) SHS capture on the BRH, BRE and BEW cDNAs. For each of the three soil cDNA samples, only the most abundant sequence clusters, encompassing ≥90% of the sequences in the H0, H1 or H2 samples, were used for the calculation. GH11 PCR sequences were clustered using a nucleotide sequence identity threshold of 95%.

A

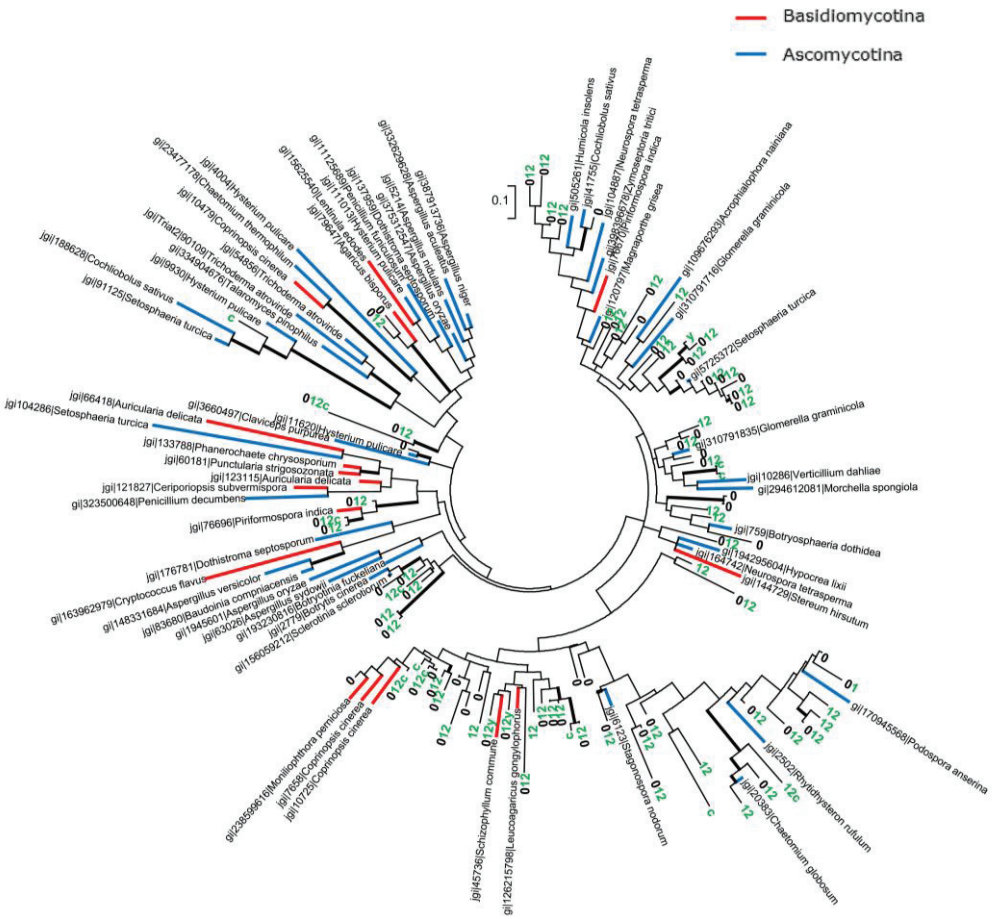

Figure S6. Phylogenetic diversity of the GH11 partial amino acid sequences obtained from (A) the PUE, (B) the BRE, (C) the BRH and (D) the BEW cDNA samples (green and black labels). 0, 1 and 2, translated PCR sequences obtained before or after one or two cycles of hybridization. Environmental cDNA sequences are scattered over the entire tree which includes representative reference sequences from Ascomycota (blue lines) and Basidiomycota (red lines). c, sequences obtained from *E. coli* clones; y: sequences functionally expressed in yeast clones. PhyML tree calculation was based on an alignment of ca. 80 amino acid long GH11 partial sequences. Thicker black branches indicate bootstrap value ≥ 60% (1000 replications).

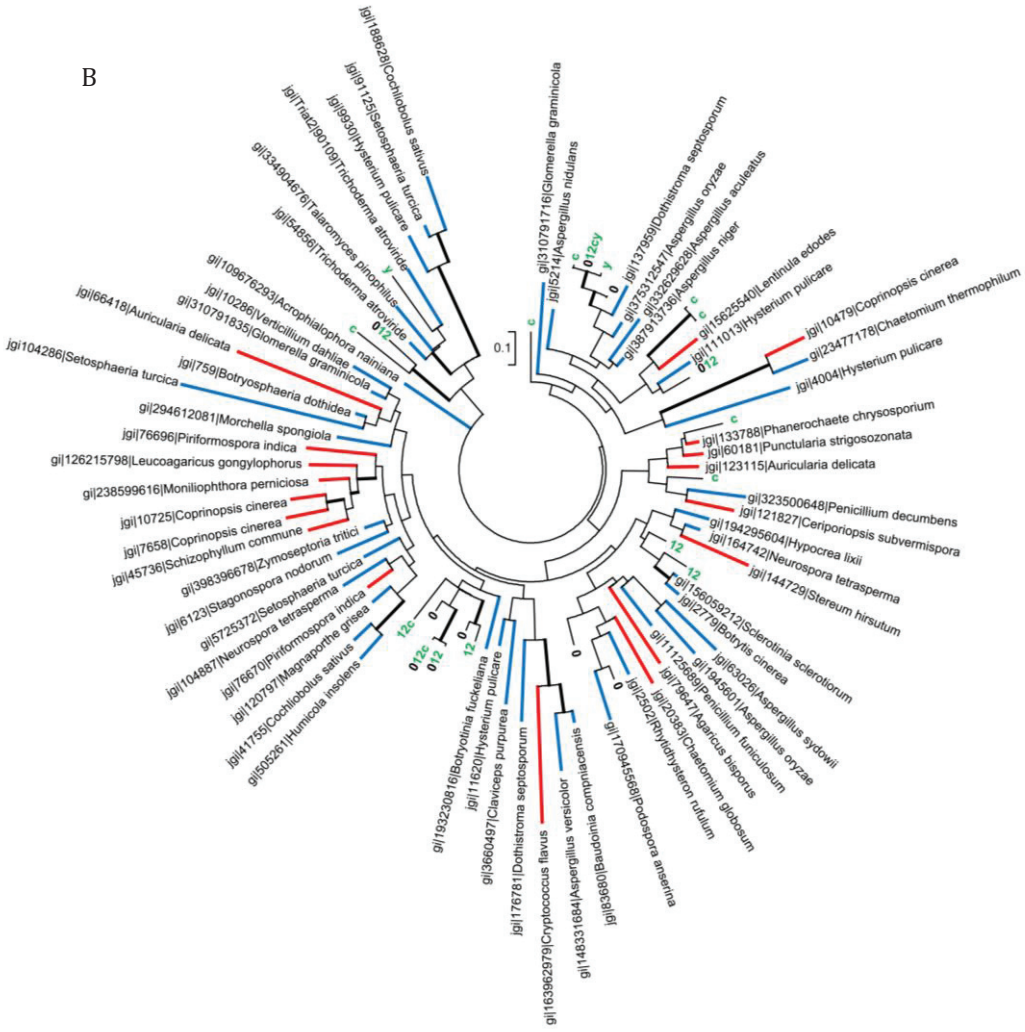

Figure S6. Phylogenetic diversity of the GH11 partial amino acid sequences obtained from (A) the PUE, (B) the BRE, (C) the BRH and (D) the BEW cDNA samples (green and black labels). 0, 1 and 2, translated PCR sequences obtained before or after one or two cycles of hybridization. Environmental cDNA sequences are scattered over the entire tree which includes representative reference sequences from Ascomycota (blue lines) and Basidiomycota (red lines). c, sequences obtained from *E. coli* clones; y: sequences functionally expressed in yeast clones. PhyML tree calculation was based on an alignment of ca. 80 amino acid long GH11 partial sequences. Thicker black branches indicate bootstrap value ≥ 60% (1000 replications).

C

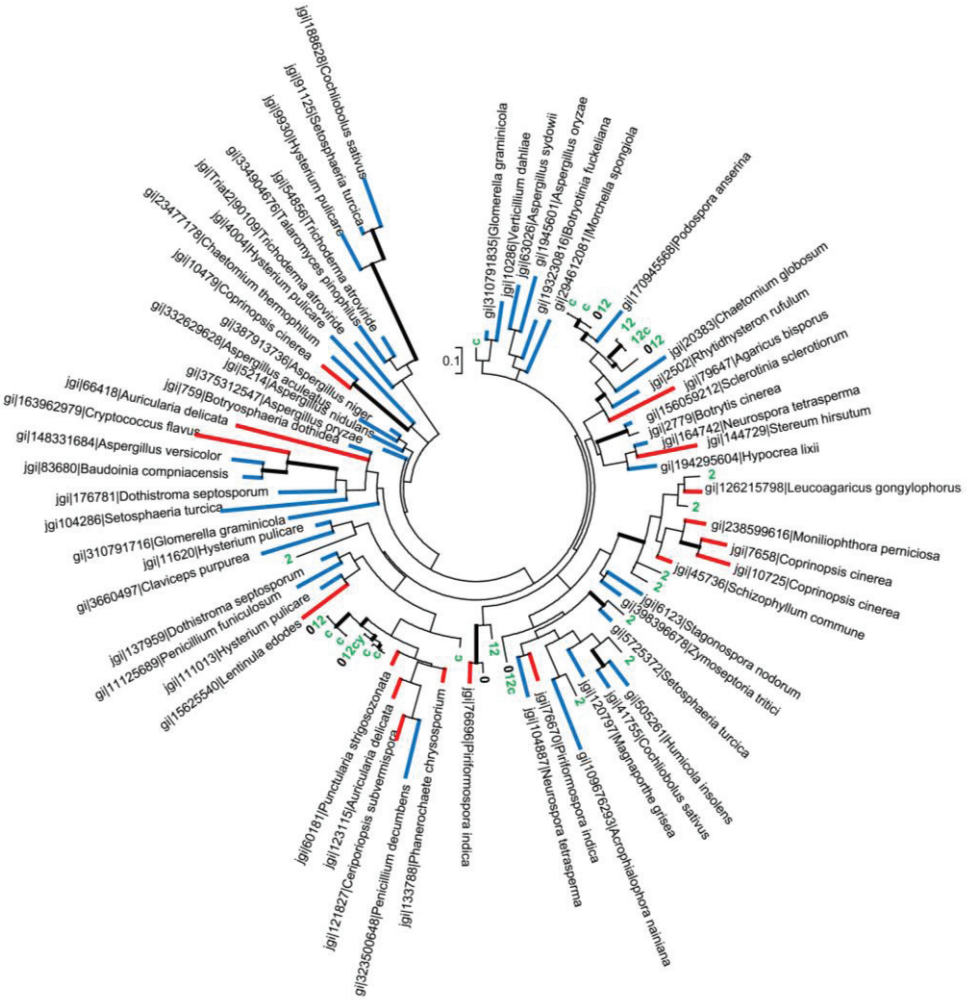

Figure S6. Phylogenetic diversity of the GH11 partial amino acid sequences obtained from (A) the PUE, (B) the BRE, (C) the BRH and (D) the BEW cDNA samples (green and black labels). 0, 1 and 2, translated PCR sequences obtained before or after one or two cycles of hybridization. Environmental cDNA sequences are scattered over the entire tree which includes representative reference sequences from Ascomycota (blue lines) and Basidiomycota (red lines). c, sequences obtained from *E. coli* clones; y: sequences functionally expressed in yeast clones. PhyML tree calculation was based on an alignment of ca. 80 amino acid long GH11 partial sequences. Thicker black branches indicate bootstrap value ≥ 60% (1000 replications).

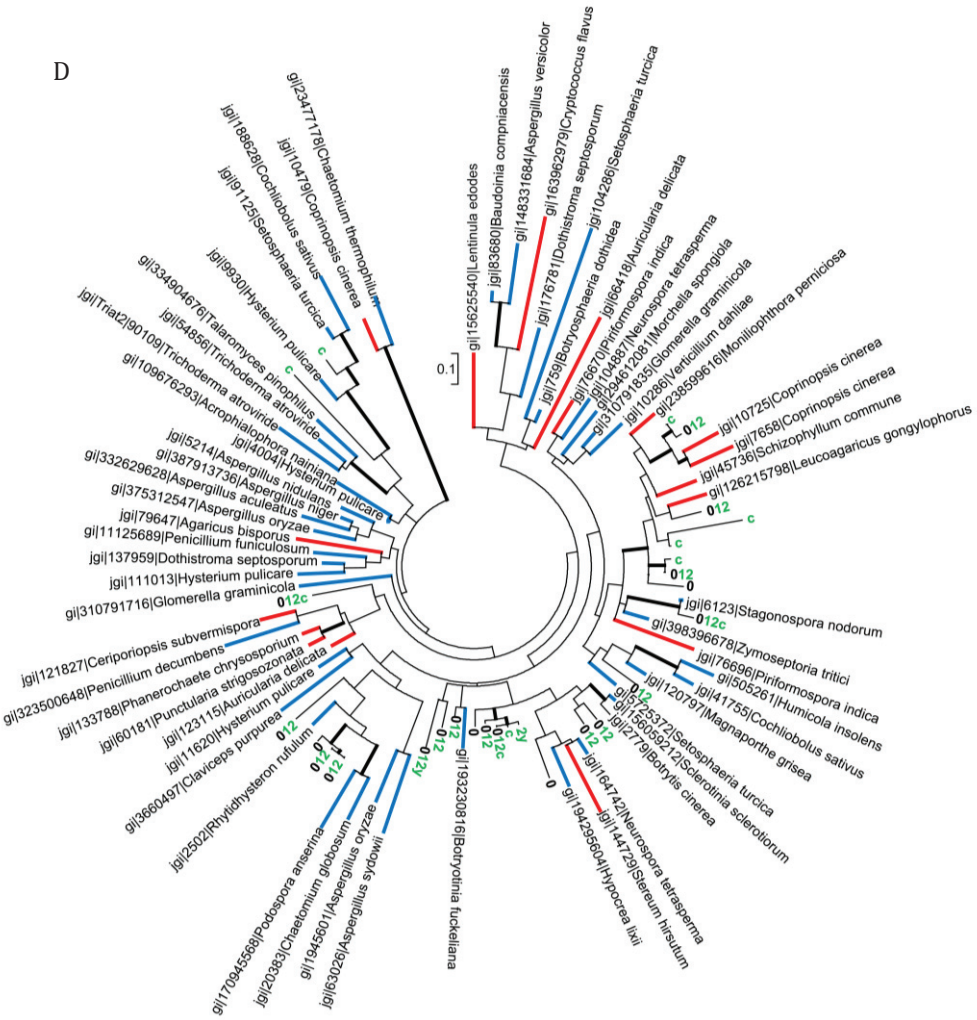

Figure S6. Phylogenetic diversity of the GH11 partial amino acid sequences obtained from (A) the PUE, (B) the BRE, (C) the BRH and (D) the BEW cDNA samples (green and black labels). 0, 1 and 2, translated PCR sequences obtained before or after one or two cycles of hybridization. Environmental cDNA sequences are scattered over the entire tree which includes representative reference sequences from Ascomycota (blue lines) and Basidiomycota (red lines). c, sequences obtained from *E. coli* clones; y: sequences functionally expressed in yeast clones. PhyML tree calculation was based on an alignment of ca. 80 amino acid long GH11 partial sequences. Thicker black branches indicate bootstrap value ≥ 60% (1000 replications).

Chapter V

PCR primers to study the diversity of expressed fungal genes encoding lignocellulolytic enzymes in soils using high-throughput sequencing

5.1 Foreword

Saprotrophic fungal communities play essential roles in plant biomass degradation in soil, a key step of carbon cycling in terrestrial ecosystems. They have developed strategies to drive plant residue mineralization by producing a large number of extracellular enzymes which hydrolyze the major components of plant litter (cellulose, hemicelluloses and lignin).

The identification of the fungal species producing these enzymes in the natural environment is essential to understand the decomposition process. Environmental genomic approaches consisting in the systematic sequencing of enzyme-coding gene families using soil-extracted RNA as material, allow for the direct *in situ* assessment of the diversity of the species involved in a specific process.

Therefore a strategy to investigate, directly in soils, the diversity of fungi performing plant biomass degradation is to use this approach by targeting expressed gene encoding lignocellulolytic enzymes. A fundamental requirement for such an approach is the design and the evaluation of gene family specific degenerated PCR primers producing fragments compatible with high- throughput sequencing approaches. These primers should be developed from conserved protein regions of each enzyme family and must be able to amplify from a wide range of genes from fungal species belonging to distant phylogenetic groups.

Within this framework, the goal of the study described in this chapter was to develop and evaluate PCR primers for the specific amplification of different fungal genes encoding enzymes belonging to different CAZyme families.

I had the opportunity to participate to this work during my stay at the University of Lyon under the supervision of Roland Marmeisse and Patricia Luis. Primers pairs designed and tested in this work were used in the two studies illustrated in *Chapters III* and *IV.*

PLOS ONE

RESEARCH ARTICLE

PCR Primers to Study the Diversity of Expressed Fungal Genes Encoding Lignocellulolytic Enzymes in Soils Using High-Throughput Sequencing

Florian Barbi[1], Claudia Bragalini[1,2], Laurent Vallon[1], Elsa Prudent[1], Audrey Dubost[1], Laurence Fraissinet-Tachet[1], Roland Marmeisse[1], Patricia Luis[1]*

1. Ecologie Microbienne, UMR CNRS 5557, USC INRA 1364, Université de Lyon, Université Lyon 1, Villeurbanne, France, **2.** Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy

*patricia.luis@univ-lyon1.fr

Citation: Barbi F, Bragalini C, Vallon L, Prudent E, Dubost A, et al. (2014) PCR Primers to Study the Diversity of Expressed Fungal Genes Encoding Lignocellulolytic Enzymes in Soils Using High-Throughput Sequencing. PLoS ONE 9(12): e116264. doi:10.1371/journal.pone.0116264

Editor: Erika Kothe, Friedrich Schiller University, Germany

Received: August 22, 2014

Accepted: November 26, 2014

Published: December 29, 2014

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All sequencing data (i.e. either generated by the Sanger sequencing or by Illumina Miseq technology) are deposited at the EMBL European Nucleotide Archive (http://www.ebi.ac.uk/ena/) under the following accession numbers: "HG799539-HG799611" and "FR875180-FR875286" for the Sanger sequences and "ERR636005-ERR636006" for the fastq-files generated by the Illumina (Miseq).

Funding: Florian Barbi was supported by a PhD grant from the French "Ministère de l'Enseignement Supérieur et de la Recherche". Claudia Bragalini was supported by a PhD grant from the University of Turin (Italy) and a CMIRA (Coopération et Mobilité Internationales Rhône-Alpes) fellowship from the French county Rhône-Alpes. The work was financed by the French Research foundation CNRS-INSU (Centre National de la Recherche Scientifique – Institut National des Sciences de l'Univers) through an "Ecosphère Continentale et Côtière" program (EC2CO MicrobiEn 2013) and an "Agence Nationale de la Recherche" project (ANR 09-GENM-033-001/ Eumetasol). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Plant biomass degradation in soil is one of the key steps of carbon cycling in terrestrial ecosystems. Fungal saprotrophic communities play an essential role in this process by producing hydrolytic enzymes active on the main components of plant organic matter. Open questions in this field regard the diversity of the species involved, the major biochemical pathways implicated and how these are affected by external factors such as litter quality or climate changes. This can be tackled by environmental genomic approaches involving the systematic sequencing of key enzyme-coding gene families using soil-extracted RNA as material. Such an approach necessitates the design and evaluation of gene family-specific PCR primers producing sequence fragments compatible with high-throughput sequencing approaches. In the present study, we developed and evaluated PCR primers for the specific amplification of fungal CAZy Glycoside Hydrolase gene families GH5 (subfamily 5) and GH11 encoding endo-β-1,4-glucanases and endo-β-1,4-xylanases respectively as well as Basidiomycota class II peroxidases, corresponding to the CAZy Auxiliary Activity family 2 (AA2), active on lignin. These primers were experimentally validated using DNA extracted from a wide range of Ascomycota and Basidiomycota species including 27 with sequenced genomes. Along with the published primers for Glycoside Hydrolase GH7 encoding enzymes active on cellulose, the newly design primers were shown to be compatible with the Illumina MiSeq sequencing technology. Sequences obtained from RNA extracted from beech or spruce forest soils showed a high diversity and were uniformly distributed in gene trees featuring the global diversity of these gene families. This high-throughput sequencing approach using several degenerate primers

constitutes a robust method, which allows the simultaneous characterization of the diversity of different fungal transcripts involved in plant organic matter degradation and may lead to the discovery of complex patterns in gene expression of soil fungal communities.

Introduction

In forest ecosystems, up to two thirds of the organic carbon (C_{org}) are stored in soils and a large part is localized in plant litters [1]. As litter input can exceed 3.5 t ha^{-1} yr^{-1}, it represents, along with root exudates, the main source of soil organic matter (OM) and its degradation by soil organisms is essential for carbon cycling [2,3]. Plant litter decomposition is largely controlled by soil fauna and microorganisms (bacteria, fungi…). In terrestrial ecosystems, the above- and belowground plant litter constitutes the main source of energy and matter for the soil heterotrophic microflora [4]. Soil microorganisms have developed strategies to drive plant-residue mineralization by producing a large number of extracellular enzymes [5].

Cellulose, hemicelluloses and lignin are the three most abundant biopolymers in plant litter and in soil OM derived from its decomposition [6]. Lignin, a polymer highly recalcitrant to enzymatic degradation, restricts microbial and enzyme access to cellulose and other labile carbon compounds that it protects [7]. Saprotrophic fungi are considered to be the most efficient decomposers of these biopolymers due to their important lignocellulolytic potential and wide diversity in soils [8,9,10]. Indeed, the complete breakdown of lignin can only be achieved by saprotrophic basidiomycetes whereas a larger number of soil fungi are able to perform the degradation of polysaccharides such as cellulose and hemicelluloses [10]. Schneider et al. [11] reported, using a metaproteomic approach, that all hydrolytic enzymes they could identify from a beech litter extract were likely of fungal origin.

Forest soil fungal communities and their associated functions can be affected by several environmental factors including seasonal climatic cycles, stand age, tree species and therefore the quality of litter they produce [12–16]. Within a given climate, litter quality is the overriding factor controlling decomposition rate [17]. Activities expressed by single fungal species are difficult to access *in situ* with conventional approaches. Methods commonly applied in soil surveys such as enzymatic activities or respiration rate measurements do not indicate which fungi are responsible for these processes [18]. Determination of the functions of fungal species, which typically requires their isolation in pure culture and the study of their effects on defined substrates, has well-documented limitations [19–21]. The use of degenerate primer sets for fungal functional gene amplification and their utilization on nucleic acids extracted from soil samples provides cultivation-independent tools for assessing the genetic diversity and activity of lignocellu-

lolytic degrading guild within fungal communities [22–27]. Relevant and unexpected information was revealed by the utilization of these approaches such as the potential role of Ascomycota species and certain ectomycorrhizal fungi in lignin degradation in soils [18, 26]. Thus far, two main fungal enzyme-coding gene families have been targeted in environmental studies; the laccases (EC 1.10.3.2) of the CAZy Auxiliary Activity (AA) family 1 whose role in lignin breakdown in the field has always been controversial [28] and the CAZy Glycoside Hydrolase (GH) family 7 encoding essentially either endo-β-1,4-glucanases (E.C. 3.2.1.4) or cellobiohydrolases (E.C. 3.2.1.176), both active on cellulose [22, 25, 29–33].

As the three major components of plant litter are cellulose, hemicelluloses and lignin, simultaneous study of fungal communities participating to their breakdown would require PCR amplification and sequencing of several key enzyme-coding gene families known to be active on these polymers. In the present study we developed, evaluated and tested on soil-extracted RNA samples PCR primers targeting functional genes active on these three plant cell wall constituents. For cellulose, in addition to the well-documented GH7 family, we developed primers for the subfamily 5 of the GH5 family (GH5-5). As the GH5 family encompasses enzymes active on a wide range of substrates (i.e. cellulose, mannane, chitosane…) which are distributed in different subfamilies (GH5-1, GH5-2, GH5-3…), we specifically targeted the subfamily GH5-5 known to comprise only endo-β-1,4-glucanases (E.C. 3.2.1.4) [34–35]. The members of this GH5-5 subfamily are widely distributed among fungi [35] and are often highly expressed in presence of cellulose in saprotrophic species [36–38]. For hemicellulose, we targeted the fungal GH11 encoding only endoxylanases and almost exclusively endo-β-1,4-xylanases (EC 3.2.1.8) [34]. Xylan constitutes the major component of hemicelluloses and its proportion is generally higher in broad-leaved trees compared to coniferous ones [39]. For lignin, we focused on the Basidiomycota class II peroxidases corresponding to CAZy Auxiliary Activity family 2 (AA2) which comprises Manganese (MnP; EC 1.11.1.13), lignin (LiP; EC 1.11.1.14), versatile (VP; EC 1.11.1.16) and generic (GP; 1.11.1.7) peroxidases. These peroxidases, which can display oxidizing activities towards aromatic compounds such as lignin, have a complex evolutionary history [40] and have been essentially studied in wood rot fungi although they have also been reported in a wide range of soil-borne Basidiomycota including symbiotic ectomycorrhizal species [24, 26–27, 41].

The present study includes different aspects. We initially designed, *in silico*, PCR primers based on sequence alignments and then tested them on a wide range of DNA samples corresponding to fungi belonging to different classes of Asco- and Basidiomycota, including species with sequenced genomes for which the presence/absence of the studied gene families had been established. We then evaluated the suitability of several of the resulting PCR products for high-throughput sequencing using the Illumina MiSeq approach. Amplifications were performed on cDNAs synthesized from soil-extracted RNA samples obtained from two forest stands differing with respect to their dominant tree species

(spruce versus beech). Working on environmental RNA, instead of DNA, gives access to the active fraction of the soil community. Furthermore, in the case of eukaryotes, it also circumvents the problem of introns, which can either disrupt PCR primer binding sites or create large size variations between PCR products.

Materials and Methods

Concerning fieldwork permits, soil samples were collected in the framework of collaborative projects including Jacques Ranger and Arnaud Legout from the "Institut National de la Recherche Agronomique" (INRA) of Nancy (France) who are in charge of the Breuil-Chenue site.

Primer design and analysis of their efficiency on fungal DNA

New degenerate primers (Table 1) were designed, according to the process described by Kellner et al. [18], to specifically amplify fungal genes encoding endo-β-1,4-glucanases (EC 3.2.1.4) of the GH5 subfamily 5, endo-β-1,4-xylanases (EC 3.2.1.8) of the GH11 family, and Basidiomycota-class II peroxidases (EC 1.11.1.13 (MnP), EC 1.11.1.14 (LiP), EC 1.11.1.16 (VP) and EC 1.11.1.7 (GP)) of the AA2 family. These primer pairs were developed on the basis of reference protein sequences retrieved from the curated CAZy [34] and GenBank [42] databases. The CAZy reference sequences (S1 Table) were compared against the NCBI database using the standard protein-protein BLAST (blastp) and the distance tree option implemented in the NCBI result page was used to display the phylogenetic relationship of each protein among the different fungal groups. Several representatives of each clade were selected to generate multiple alignments. Degenerate primer pairs were then developed for conserved protein regions of each enzyme family or subfamily (S1 Table) to amplify either both Ascomycota or Basidiomycota (GH5-5 and GH11) or only Basidiomycota (AA2).

The efficiency of each primer pair was tested on DNA extracted from 72 different fungal species, belonging to either the Basidiomycota or Ascomycota (S2 Table). Complete genomes were available for 27 of these fungal species at the JGI genome portal MycoCosm (http://genome.jgi-psf.org/programs/fungi/index.jsf; [43]) or at the Broad Institute website (https://www.broadinstitute.org). Fungal genomic DNA extraction was performed from mycelia or fresh fruit bodies as previously described [44]. For PCR amplifications, 60 ng of fungal DNA were added to a 25 µl reaction mixture containing 2.5 µl of 10X polymerase buffer (Invitrogen), 0.75 µl of $MgCl_2$ (50 mM), 2.5 µl of dNTPs (2 mM each), 1 µl of 5 $mg.ml^{-1}$ bovine serum albumin, 0.5 µl of each primer (20 µM) and 0.1 U of polymerase mix (1:24 (U:U) of Biorad iProof High-Fidelity DNA Polymerase : Invitrogen *Taq* DNA polymerase). Cycling conditions, performed on a Peltier Thermal Cycler 100 (MJ Research), were 94°C for 3 min; 45 cycles of 94°C for 45 s, 50°C for 45 s, 72°C for 45 s, followed by 72°C for 10 min. Control reactions without nucleic acid were always run in parallel. PCR products were extracted

Table 1. Degenerate primers used in this study.

Gene family-Subfamily[1]	Primer names and sequences (5′-3′)	PCR fragment length (bp)	Targeted fungal groups	Reference
GH7	fungcbhI-F: ACC AAY TGC TAY ACI RGY AA; fungcbhI-R: GCY TCC CAI ATR TCC ATC	515	Basidiomycota; Ascomycota	[23]
GH5-5	fungGH5-5-F: GAR ATG CAY CAR TAC CTY GA; fungGH5-5-R: CA NGG ICC RGC RGC CCA CCA	248	Basidiomycota; Ascomycota	This study
GH11	fungGH11-F: GGV AAG GGI TGG AAY CCN GG; fungGH11-R: TG KCG RAC IGA CCA RTA YTG	281	Basidiomycota; Ascomycota	This study
AA2	basidioAA2-F: GGY GGI GGI GCB GAY GGY TC; basidioAA2-R: GG RGT IGA GTC RAA NGG	398	Basidiomycota	This study

[1]according to the CAZy database (http://www.cazy.org/; [34]). The GH7 family encodes essentially either fungal endo-β-1,4-glucanases (E.C. 3.2.1.4) or cellobiohydrolases (E.C. 3.2.1.176), both active on cellulose; the GH5-5 subfamily only encodes fungal endo-β-1,4-glucanases (E.C. 3.2.1.4); the GH11 family encodes only endoxylanases and almost exclusively fungal endo-β-1,4-xylanases (EC 3.2.1.8); The AA2 family comprises Manganese (MnP; EC 1.11.1.13), lignin (LiP; EC 1.11.1.14), versatile (VP; EC 1.11.1.16) and generic peroxidases (GP; 1.11.1.7).

doi:10.1371/journal.pone.0116264.t001

from agarose gel using the QIAquick Gel Extraction Kit (Qiagen), ligated to the plasmid pCR4-TOPO (TOPO TA Cloning kit for sequencing, Invitrogen) and introduced into chemically competent TOP 10 *E. coli* cells. After plasmid extraction, using the NucleoSpin Plasmid kit (Macherey Nagel), inserts were sequenced by Biofidal (Lyon, France) using M13 forward and reverse primers and the Sanger technology.

Study site and soil sampling

The experimental site of Breuil-Chenue forest is located in the Morvan mountains, Burgundy, France (47°18′10″N, 4°4′44″E). The elevation is 640 m, mean annual rainfall is 1280 mm and mean annual temperature is 9°C. The parent rock is granite and the soil is an alocrisol with pH between 4.0 and 4.5 [45]. The site is an environmental research observatory set up in 1976 in order to study the effects of tree species substitution on the biochemical and biological functioning of the soil ecosystem. The original forest, composed of several broadleaved tree species was clear-cut and replaced by several mono-specific stands of either coniferous or broadleaved tree species. Soil samples were collected from beech (*Fagus sylvatica*) and spruce (*Picea abies*) stands (ca 1000-m^2 plots with soil pH of 3.9 and mull humus layers) along a systematic sampling grid [46]. For each stand, 14 soil cores of 750 cm^3 (8 cm in diameter, 15 cm in depth) were sampled in one plot in July 2007 and 14 soil cores were taken in three independent plots in July and October 2010. After removing the surface litter, the organic matter-rich horizons (depth ± 0–7 cm) of each soil core were homogenized separately and sieved (2 mm mesh size) to eliminate small debris and root fragments. For each sampling date, a single composite sample per plot was prepared by pooling 100 mL of each core. Subsamples of the composite samples were frozen and kept at −70°C.

Soil RNA extraction, reverse transcription and PCR amplification

Total RNA was extracted as described in Damon et al. [46] from 78 to 104 g of soil for each composite sample. Three to four series of 40 extractions (0.65 g of soil each) were performed in parallel for each composite samples. RNA extracts obtained from soil samples collected in July and October 2010 were pooled by forest stand before reverse transcription. Samples were designated as BB2007 for Breuil Beech 2007, BS2007 for Breuil Spruce 2007, BB2010 for Breuil Beech 2010 and BS2010 for Breuil Spruce 2010. Double stranded cDNA was synthesized from 2 µg of total soil RNA using the Mint-2 cDNA synthesis kit (Evrogen). The optimal number of PCR cycles to maintain a balance between maintaining transcript representation and reducing nonspecific background amplification during the cDNA production was found to be 27 cycles. The resulting cDNAs were used as template to specifically amplify expressed fungal genes encoding endo-β-1,4-glucanases (GH5-5 subfamily), endo-β-1,4-xylanases (GH11 family), and Basidiomycota-class II peroxidases (AA2 family) by using primers developed in this study and cellulases (GH7 family) by using the degenerate primers designed by Edwards et al. [23]. All PCR amplifications were performed in quintuplet in 25 µl reaction mixtures containing 1 µl of cDNA, 2.5 µl of 10X polymerase buffer with 25 mM $MgCl_2$ (Invitrogen), 2.5 µl of dNTPs (2 mM each), 0.5 µl of each primer (20 µM) and 0.1 µl polymerase mix (1:24 (U:U) of Biorad iProof High-Fidelity DNA Polymerase : Invitrogen *Taq* DNA polymerase). Amplification conditions were 3 min at 94℃; either 35 or 45 cycles of 45 s at 94℃, 45 s at either 48℃ for the GH7 primer set or 50℃ for all other primers, 2 min at 72℃, followed by 10 min at 72℃. Control reactions with non-reverse transcribed mRNA and without nucleic acid were run in parallel.

Amplicon sequencing by Sanger and Illumina MiSeq approaches

Amplicons from samples BB2007 and BS2007 were sequenced using a Sanger approach. PCR products were extracted from agarose gel and cloned in *E. coli* as described above. A total of 48 bacterial clones were randomly selected for each gene family and the inserts sequenced by AGOWA (Germany) using both M13 forward and reverse primers. PCR products obtained from samples BB2010 and BS2010 were subjected to an Illumina MiSeq sequencing. Amplicons from the five independent PCR reactions were pooled and directly purified using the Qiagen QIAquick PCR purification Kit (GH11 and GH7 families) or separated by electrophoresis and gel purified (AA2 family). DNA quantity was measured using a Qubit 2.0 fluorometer (Invitrogen) and the Qubit dsDNA HS assay kit (Invitrogen). For each sample (BB2010 and BS2010) an equimolar mix of the different PCR products was made and a paired-end sequencing was carried out on an Illumina MiSeq sequencer by FASTERIS (Switzerland) using a 2×250 bp sequencing kit as available at that time. GH5-5 amplicons could not be included in this MiSeq analysis as their sizes (<250 bp) were too divergent from the ones of the class II peroxidases (~400 bp) and GH7 (~515 bp).

Sequence analysis

All Sanger sequences obtained from fungal DNA or soil cDNA were manually edited, corrected, and deposited at the EMBL European Nucleotide Archive (http://www.ebi.ac.uk/ena/) under accession numbers HG799539-HG799611 and FR875180-FR875286. Concerning the Miseq data (deposited at EMBL-ENA under the project number PRJEB7363 with fastq-files accession nos. ERR636005-ERR636006), paired-end reads were assembled using PandaSeq v.2.5 [47] and all sequences containing "N"s were filtered out. Assembled paired-end reads were then analyzed using Mothur v.1.33 [48]. As GH11 and AA2 amplicons were smaller than 400 bp, forward and reverse reads were confidently assembled and the resulting merged sequences were trimmed according to both primer sequences using Mothur. As the GH7 PCR products were of about 500 bp in length, forward and reverse reads could not be assembled and we limited our analysis to the first 210 nucleotides of the reads bordered by the GH7 forward primer. To ensure high quality data for analysis, assembled pair-reads containing homopolymers longer than 7 bp and more than two mismatches in any primer sequence were removed. Chimeric sequences were detected using the UCHIME algorithm [49] and removed from the datasets. Sequences were then clustered at a cutoff value of 95% sequence identity for the GH7 & GH11 families and 93% for the AA2 one. For each sequence cluster, the most abundant sequence was chosen as its representative.

The subfamily assignation of the partial GH5 sequences was performed by Dr. Bernard Henrissat (CNRS, Marseille) using Hidden Markov models specific for each of the GH5 subfamilies and that are used for the daily updates of the GH5 subfamily information in the CAZy database. The subfamily assignation of the partial AA2 was performed as described in Kellner et al. [27] by recording the presence or absence of specific amino acids in the protein sequences. All MnP sequences contain an aspartic acid residue, corresponding to the Asp-175 of the *Phanerochaete chrysosporium* MnP1 (AAA33744), which is crucial for Mn^{2+} oxidation. All LiP sequences display a tryptophan residue, equivalent to the Trp-171 of the *P. chrysosporium* LiPH8 (AAA53109), which is responsible for oxidation of phenolic compounds. All VP sequences possess these two amino acids, while GP sequences lack them both.

Diversity indices and richness estimators were calculated using EstimateS v.9.1.0. (http://viceroy.eeb.uconn.edu/estimates/) on subsamples containing the same number of sequences per sample (i.e. the sequence number obtained for the BB2010 sample) to eliminate the effect of sequencing effort.

Phylogenetic analyses were performed on deduced amino acid sequences, which were aligned using MUSCLE [50] to all homologous sequences retrieved from GenBank (http://www.ncbi.nlm.nih.gov/; [42]), CAZy (http://www.cazy.org/; [34]), PeroxiBase (http://peroxibase.toulouse.inra.fr/; [51]) databases and from the published fungal genome sequences available at the JGI and Broad Institute. Maximum-likelihood (ML) trees were generated with PhyML v.3.0 [52] using the WAG substitution model [53] as implemented in SeaView v.4 [54]. Robustness of

the tree topology was tested by bootstrap analysis (1000 replicates). As there were too many homologous protein sequences in databases (i.e. more than 500 different sequences for the AA2 and GH7 families), the ML trees illustrated in the manuscript include ~100 known protein sequences representative of the phylogenetic diversity of each gene family, including the ones generated in the present study. Moreover, all environmental sequence clusters, except singletons identified in only one of the two soil samples, were included in these phylogenetic trees drawn in FigTree v.1.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Results

Primer specificity and efficiency on fungal DNA and soil cDNA

In silico-designed primers for fungal genes belonging to families GH5 (subfamily 5), GH11 and AA2 were first tested on DNA extracted from known fungal species belonging to either the Ascomycota or the Basidiomycota (S2 Table). Depending on the gene family or subfamily, PCR products of the expected size (i.e. ~250 bp for GH5-5, ~300 bp for GH11 and ~400 bp for AA2 family) were obtained for 17 to 47 of the 72 fungal species tested (S2 Table). The sequencing of two cloned PCR products per gene for 28 of these 72 fungal species always gave sequences belonging to the expected gene family or subfamily (GH5-5, GH11 or AA2), thus demonstrating the specificity of the PCR primers. Furthermore, in 22% of the cases the two homologous sequences were different, as expected for genes often occurring as gene families in fungal genomes. All GH5 partial sequences obtained were assigned to the GH5-5 subfamily encoding only endo-β-1,4-glucanases. Concerning the AA2 family, all subfamilies could be amplified. Indeed, among the 13 different AA2 sequences obtained, 8 belonged to the MnP subfamily, 1 to the LiP subfamily, 2 to the VP family and 2 to the GP one (S2 Table).

As anticipated, both GH5-5 and GH11 primers, which were designed using both Ascomycota and Basidiomycota gene sequences, indifferently amplified genes from DNA extracted from either Ascomycota or Basidiomycota species distributed in different families, orders and classes of these taxonomic groups. On the contrary, AA2 primers, designed using sequences exclusively from Basidiomycota amplified sequences only from DNA extracted from species belonging to this taxon (S2 Table).

The effectiveness of the degenerate primers could also be evaluated from the amplification results obtained using DNA extracted from 27 species whose genome sequences, and therefore the presence or absence of the studied gene families, were known. Depending upon the gene family, positive amplifications were ranged from for 69% (9 of the 13 fungal species with predicted GH11 genes in their genomes) to 85% (23 of the 27 fungal species with predicted GH5 genes in their genomes) of the expected cases (Fig. 1). Positive amplifications were never observed when the genes were known to be absent. In cases of apparent amplification failure, examination of the genomic copie(s) of the corresponding gene families indicated that absence of amplification resulted essentially from

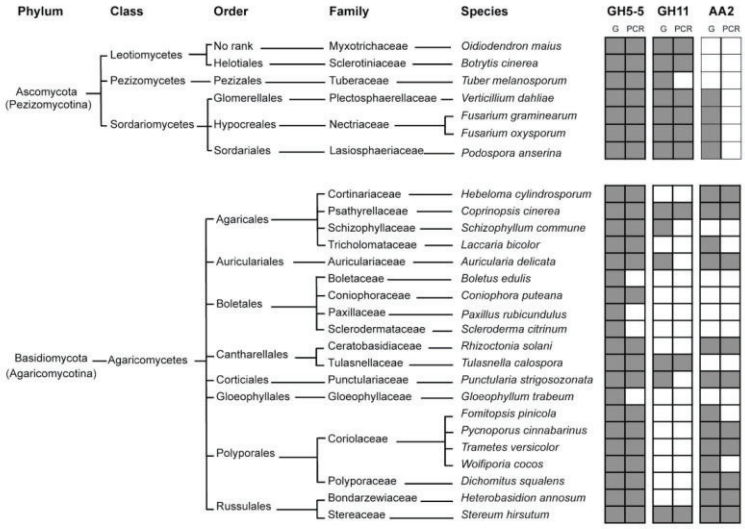

Fig. 1. Primer specificity. Each pair of degenerate primers was tested on DNA extracted from 27 sequenced fungal species. Presence or absence of the GH5 (subfamily 5), GH11 and AA2 gene families in the published genomes (G) are indicated by gray and white boxes, respectively. Similarly, positive or negative PCR amplifications (PCR) are materialized by grey and white boxes respectively.

doi:10.1371/journal.pone.0116264.g001

highly divergent sequences at the PCR primer binding site(s) and to a lesser extent from the presence of introns within the primer region. Furthermore, for the same reasons, not all predicted copies in genomes are amplifiable with the designed primer sets. Concerning fungal species with predicted GH5-5 genes in their genomes, 33 to 100% of the genomic copies are amplifiable depending on the species (S2 Table). Regarding the GH11 genes, for most species, 33 to 50% of the predicted genomic copies are amplifiable. Based on the total number of predicted Class-II peroxidase sequences within the 21 Basidiomycota-genomes, our AA2 primers preferentially target MnP and LiP subfamilies and to a lesser extend GP and VP subfamilies as they potentially amplify 79% of LiP, 49% of MnP and only 12% of GP or VP genomic copies (S2 Table).

As the total RNA extracted from soil samples contains mRNA expressed by all soil organisms and not only by fungi, the specificity and efficiency of the designed degenerate primers (for the GH5-5 subfamily, GH11 and AA2 families) were also evaluated on soil cDNA by the sequencing of cloned PCR products. This analysis was carried out on the cDNA obtained from soil samples collected in 2007 in a beech (sample BB2007) and a spruce (sample BS2007) forest stand. For each cDNA sample, a total of 48 sequences were analyzed for each gene family or

subfamily (i.e. GH5-5, GH7, GH11 and AA2). Among these sequences, between 8% and 29% of them were of bad quality and the presence of chimeric sequences was detected for the AA2 gene fragments from the beech forest soil (S3 Table). Analysis of the remaining high-quality sequences showed that all of them were homologous (with percentages of similar amino acid positions above 70%) to fungal lignocellulolytic enzyme sequences of the corresponding families/subfamilies already deposited in public databases (S3 Table). All GH5 partial sequences were assigned to the GH5-5 subfamily. Concerning the 23 different Basidiomycota AA2-sequence clusters obtained, 16 corresponded to MnP, 4 to GP and 3 to LiP. No VP sequence was detected. For each soil sample, the number of sequence clusters detected with the newly degenerate primers for families GH5 (subfamily 5), GH11 and AA2 was similar or higher to the number of sequence clusters obtained with the already published and widely used GH7 primers (S3 Table).

MiSeq amplicon sequencing of multiple fungal lignocellulolytic gene family from soil cDNA samples

For the Illumina MiSeq sequencing, GH7, GH11 and AA2 PCR products obtained from one cDNA sample were pooled prior sequencing and a total of 12934 and 5761 assembled pair-reads (without "N"s) were obtained for the BS2010 and BB2010 soil samples, respectively. To ensure the highest quality of the datasets we removed assembled pair-reads containing primer mismatches or homopolymers and all suspected chimeric sequences to retain a total of 8910 (BS2010) and 2425 (BB2010) sequences for further analyses (Table 2). To determine at which percentage of identity sequences should be clustered, we first plotted for each gene family the number of sequence clusters against the percentage of identity used for clustering and then calculated the differences in the number of sequence clusters at n+1% identity versus the one at n% (S1 Fig.). The cutoff value was systematically set at the beginning of the plateau when the difference in the number of sequence clusters between two successive percentages of dissimilarity represented less than 6% of the number of clusters obtained at the cutoff value of 0% dissimilarity (S1 Fig.). As a result, sequences were clustered at a similarity cutoff value of 95% for the GH11 (S1-A Fig.) and GH7 sequences (S1-B Fig.) and of 93% for the AA2 ones (S1-C Fig.).

Altogether, between 19 and 115 different sequence clusters were detected within the initial datasets (Table 2). Clusters containing a single sequence (i.e. so-called singletons) represented between 48 and 68% of these sequence clusters (Table 2). As 20% of these singletons contained at least one "stop codon" in the predicted ORF (data not shown), all unique sequences detected in only one soil sample were assumed to result from sequencing errors and removed from the initial datasets before further analyses. To control that removing singletons from the initial datasets did not modify their diversity patterns, we computed the Shannon indexes of each dataset including or not these singletons (Table 2). For each gene family, each dataset was first rarefied to the same sequence depth (i.e. the lowest

Table 2. Illumina MiSeq sequencing results obtained for Breuil Spruce (BS2010) and Breuil Beech (BB2010) forest soils.

	BS2010 GH7	BB2010 GH7	BS2010 GH11	BB2010 GH11	BS2010 AA2	BB2010 AA2
Total no. of sequences	1068	490	6060	1382	1782	553
No. of clusters including singletons	115	38	84	23	102	19
No. of singletons (%)	67 (58%)	26 (68%)	53 (63%)	11 (48%)	60 (59%)	10 (53%)
No. of sequences used in subsamples with (without) singletons	490 (465)	490 (465)	1382 (1372)	1382 (1372)	553 (542)	553 (542)
No. of clusters in subsamples with (without) singletons	83 (41)	38 (13)	40 (24)	23 (13)	49 (39)	19 (9)
Shannon index with (without) singletons[1]	3.34 (3.01)	1.87 (1.58)	2.38 (2.27)	1.43 (1.38)	2.83 (2.77)	1.69 (1.59)
S_{chao1} with (without) singletons[1]	251 (44)	103 (13)	82 (31)	41 (13)	82 (48)	64 (9)

Sequence statistics, diversity indexes and richness estimators calculated on subsamples generated from the initial sequence datasets (including singletons) or from datasets without the singletons.
All functional genes encode lignocellulolytic enzymes, i.e. fungal cellulases (GH7), fungal endo-β-1,4-xylanases (GH11), and Basidiomycota class-II peroxidases (AA2).
[1]Shannon indexes and S_{chao1} estimators were calculated using datasets, which contained identical number of sequences for both forest soils.

doi:10.1371/journal.pone.0116264.t002

sequence value between the two forest soils) before index calculation to eliminate the effect of sequencing effort. Similar Shannon indexes were observed for all datasets irrespective of the presence or absence of singletons (Table 2).

Higher gene diversity was systematically observed for the spruce forest stand. In the case of the AA2 gene family, the predicted number of sequence clusters (S_{Chao1} estimator) for the AA2 family was estimated at 48 for the spruce soil sample (BS2010) and only at 9 for the beech soil sample (BB2010). As illustrated in Fig. 2, ~16 to 34% of the non-singleton sequence clusters contained 75% of the sequences. The percentage of shared sequence clusters between the two forest soils was low as it represented between 2 to 10% of the total number of clusters; which corresponded to only between 1 and 24% of the total number of sequences.

Phylogenetic analysis of the environmental sequences

Deduced amino-acid sequences from the soil cDNA sequences obtained by either Sanger or MiSeq sequencing were used for phylogenetic analyses. For the MiSeq sequences, the analysis was limited to the non-singleton sequence clusters of families GH7, GH11 and AA2. In addition to these environmental sequences, we also included in the alignments sequences obtained in the present study from fungal-extracted DNA as well as published fungal sequences representative of the diversity of each gene family. For both the AA2 (Fig. 3) family and GH5-5 (S2 Fig.) subfamily, a clear separation between Ascomycota and Basidiomycota sequences was observed in phylogenetic analyses, thus allowing confident assignation of anonymous environmental sequences to these taxa. This was clearly not the case for the GH11 (Fig. 4) and GH7 (S3 Fig.) gene families for which sequences from Ascomycota and Basidiomycota intermingled in the phylogenetic trees. Sequences, amplified using the newly designed GH5-5, GH11 and AA2 primers (from either fungal DNA or soil cDNA) were distributed over the entire

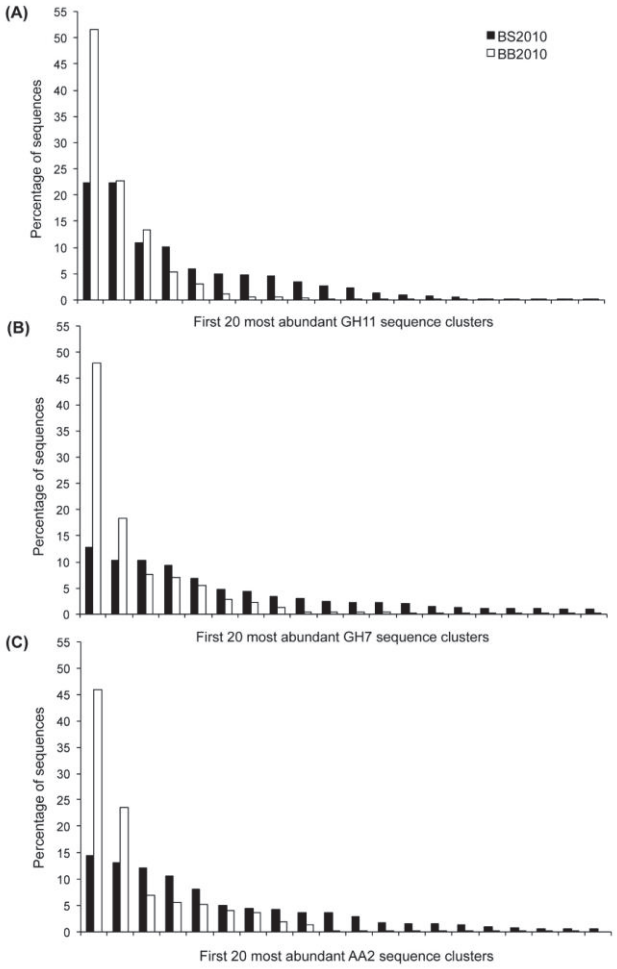

Fig. 2. Distribution of the Illumina MiSeq sequences within the 20 first most abundant GH11 (A), GH7 (B) and AA2 (C) sequence clusters. Nucleotide sequences obtained from the two studied forest soils collected in 2010 under spruce (BS2010) and beech (BB2010) were clustered at 95 (GH11 and GH7) or 93% (AA2) identity threshold.

doi:10.1371/journal.pone.0116264.g002

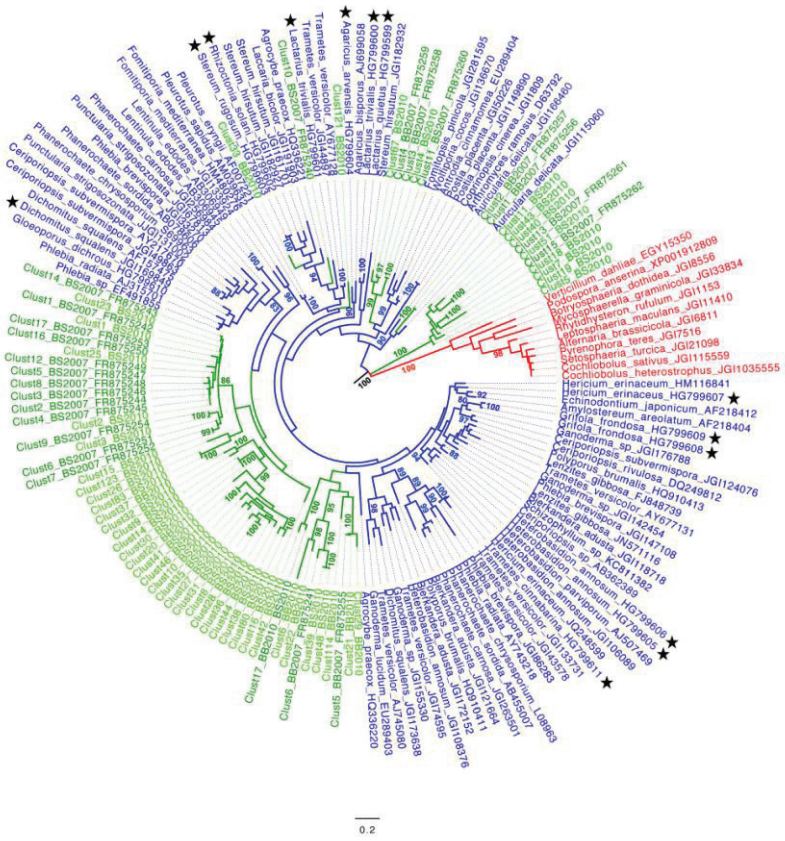

Fig. 3. Phylogenetic positions of Basidiomycota class II peroxidase (AA2) amino-acid sequences. Basidiomycota AA2 amino-acid sequences were deduced from the nucleotide sequences amplified from beech (BB) or spruce (BS) soil cDNAs. The Maximum-likelihood phylogenetic tree includes all Sanger sequences amplified from the 2007 soil cDNA samples (BS2007 and BB2007) and all non-singleton sequence clusters detected by Illumina MiSeq sequencing of the 2010 soil cDNA samples (BS2010 and BB2010). Representative Ascomycota and Basidiomycota sequences are marked in red and blue, respectively, whereas the environmental sequences appear in green. Stars identify reference sequences obtained in the present study. Robustness of the tree topology was tested by bootstrap analysis (1000 replicates) and only bootstrap values ≥80 are given.

doi:10.1371/journal.pone.0116264.g003

corresponding gene trees. However, as manifest in the AA2 (Fig. 3) and GH11 (Fig. 4) phylogenetic gene trees, the environmental sequences tend to group together. Only few environmental sequences were found closely related to known

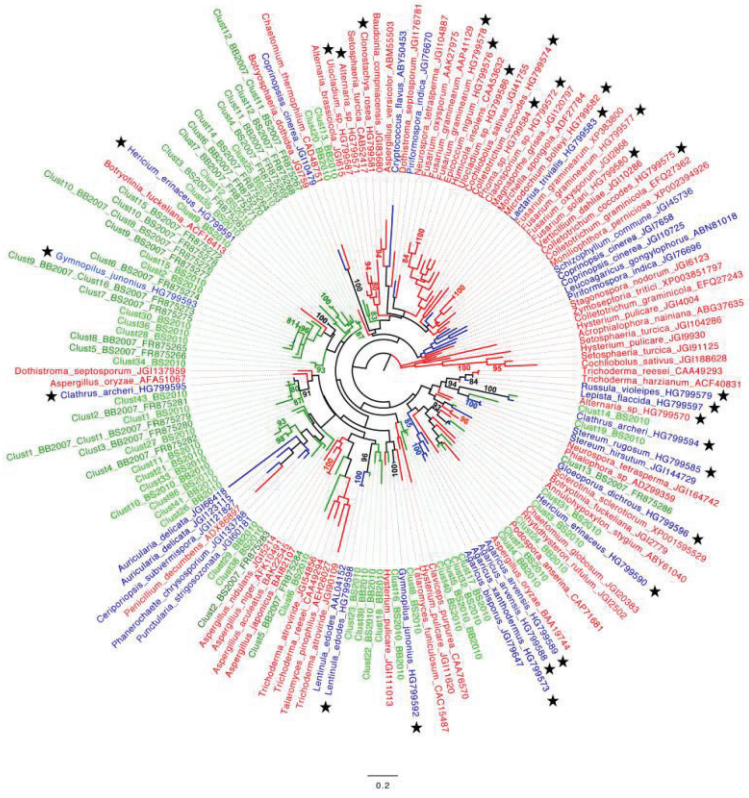

Fig. 4. Phylogenetic positions of fungal endo-β-1,4-xylanase (GH11) amino-acid sequences. Fungal GH11 amino-acid sequences were deduced from the nucleotide sequences amplified from beech (BB) or spruce (BS) soil cDNAs. The Maximum-likelihood phylogenetic tree includes all Sanger sequences amplified from the 2007 soil cDNA samples (BS2007 and BB2007) and all non-singleton sequence clusters detected by Illumina MiSeq sequencing of the 2010 soil cDNA samples (BS2010 and BB2010). Representative Ascomycota and Basidiomycota sequences are marked in red and blue, respectively, whereas the environmental sequences appear in green. Stars identify reference sequences obtained in the present study. Robustness of the tree topology was tested by bootstrap analysis (1000 replicates) and only bootstrap values ≥80 are given.

doi:10.1371/journal.pone.0116264.g004

reference ones. Few environmental sequences (i.e. 6 for the AA2, 5 for the GH5-5 and 3 for the GH7 family) detected in the soil samples collected in 2007 were also found in the soil samples collected in 2010.

Discussion

Decomposition of plant organic matter is an enzymatically complex process largely mediated by consortia of fungal species that act simultaneously or successively [15, 55] and which need to be followed as they can be affected by several environmental factors [16, 56–57].

In the present study, we designed three pairs of degenerate primers to specifically amplify fungal lignocellulolytic genes (belonging to the GH5-5 subfamily, GH11 and AA2 families) involved in the hydrolysis of complex plant polymers (cellulose, hemicelluloses and lignins respectively) and used two of them (GH11 and AA2) along with the available GH7 primers [23] in a preliminary high-throughput MiSeq sequencing. These new primers were designed to generate fragments compatible with the Illumina MiSeq approach (i.e. PCR fragments smaller than 400 bp). As the GH5 family encompasses several catalytic activities, as opposed to Kellner et al. [18], we specifically targeted the fungal subfamily GH5-5 known to only encode endo-β-1,4-glucanases active on cellulose [34–35]. Moreover, the Basidiomycota-AA2 primers were designed to potentially amplify all class II peroxidases subfamilies (i.e. MnP, LiP, VP and GP) and not only the MnP one. The increasing number of fungal genomes and their taxonomic breadth [43] allowed us to improve the effectiveness and the specificity of the newly designed primers. Effectiveness of these degenerate primers was demonstrated by the relatively high correspondence (above 69%) between the known presence of these genes in 27 sequenced fungal genomes and their positive amplification from the DNA extracted from these species (Fig. 1). Effectiveness of these primers is also apparent from the broad distribution of the sequences, amplified either from fungal DNA or soil cDNA, across the different corresponding gene trees, which summarize the phylogenetic diversity of each gene family. Full primer universality is however impossible to achieve as experimentally demonstrated by e.g. Hadziavdic et al. [58] for the non-coding 18S rRNA gene widely used in metabarcoding of eukaryotic microbial communities. The situation is even worse for protein-coding gene families, which show better conservation at the amino-acid level than at the nucleotide one due to the degeneracy of the genetic code. Protein-coding gene families, such as those coding for lignocellulolytic enzymes also display complex evolutionary histories leading to the presence of highly variable numbers and sometimes unrelated homologous genes within and between fungal genomes [40, 59–60]. As a consequence the design of degenerate primers allowing the amplification of all gene copies from all species represents an unachievable aim. Based on the sequenced fungal genomes, we estimated that 33 to 100% of the gene copies were amplifiable with the GH5-5 designed primers and 33 to 50% for the GH11 ones. Concerning the AA2 family, if all subfamilies were amplified, our primers seemed to preferentially target LiP and MnP encoding genes (S2 Table). Specificity of the designed degenerate primers was confirmed as only the targeted genes were amplified from soil cDNA (results from either Sanger or Illumina MiSeq sequencing).

A majority of recent studies using degenerate primers targeted a single expressed functional gene family and sequenced amplicons using either the Sanger or the 454 pyrosequencing approach [26–27, 32]. Due to (i) the increasing length of the sequenced fragments (at present 2×300 bp), (ii) the very high output (25 million of reads per run) and (iii) the fixed length of the reads generated by the Illumina MiSeq platform, this NGS technology is replacing the pyrosequencing as the method of choice for low-cost and high-quality sequencing [61]. As such, the present study is the first one, which evaluates simultaneously the diversity of transcribed fungal genes encoding different enzymes active on plant cell wall polymers (cellulose, hemicelluloses and lignins) using the Illumina MiSeq technology. The matrix was soil RNA which contains a low proportion of mRNA, estimated at less than 10% by Urich et al. [62] and far less if we only consider eukaryotic mRNA [63]. Among these mRNA, fungal transcripts encoding specific categories of lignocellulolytic enzymes represent themselves a small proportion that has been estimated by systematic sequencing of forest soil eukaryotic cDNAs [64]. Among c.a. 16,000 cDNAs from the same spruce and beech soils used in the present study, only between two to seven transcripts corresponding to CAZy families GH5, GH7, GH11 and AA2 were identified [64]. Therefore, amplification of targeted sequences with degenerate primers combined with high-throughput sequencing certainly represents the most straightforward way to assess the diversity of specific functional gene categories in soils [33].

To analyze the diversity of any environmental sequence dataset obtained by metabarcoding, two parameters must be evaluated; (i) at which percentage of identity should be clustered the sequences and (ii) what is the biological significance of singletons and should they be taken into consideration? Regarding sequence clustering, it cannot be done at a fixed cutoff for all gene families as different genes evolve at different evolutionary rates [65]. Fungal lignocellulolytic gene families have complex evolutionary histories (characterized by multiple independent gene loss/acquisition events affecting homologous copies) leading to the presence of one or several, either highly similar or divergent, copies per genome [40, 59–60, 66]. As a consequence we empirically defined a different cutoff for each of the gene families as described in the result section. These cutoffs of 95% identity between DNA sequences for the GH7 & GH11 sequences and of 93% for the AA2 ones are somehow lower compared to those usually adopted in metabarcoding for non-translated rRNA sequences (usually $\geq$97%; [67]). This may reflect the coding nature of the corresponding sequences and the associated degeneracy of the genetic code. Interestingly, by clustering the GH7 sequences at a 95% identity threshold, we predicted a number of sequence clusters for the BS2010 spruce soil of 44 (Chao1 richness estimator, Table 2), a value similar to the 46 ± 9 GH7 clusters per sample estimated by Baldrian et al. [33] for the humic horizon of another spruce forest soil.

Concerning the ecological significance of singletons, as about 20% of them contained at least one "stop codon" in the predicted ORF, we assumed that a majority of them may have arisen from sequencing errors. Moreover, as for other NGS datasets (e.g. [68]), Shannon diversity indices calculated for each functional

gene family after rarefying the datasets from both forest soils to the same sequencing depth showed similar values independently of the presence or absence of these singletons (Table 2). We therefore opted for not taking into account these sequences. Furthermore, considering the typical distribution of fungal-taxa abundances in soils where only few taxa are highly abundant [67–69], for the low abundance taxa, even if they are highly active, their sequences will be likely often retrieved as singletons. The biological/technical significance of singletons is thus questionable and will always be affected by sequencing depth and errors [70].

As for the assignation of soil functional gene sequences to fungal taxa, it presently suffers from a lack of sequence information in public databases [22–23, 27, 32] despite recent efforts to sequence the genomes of fungal species representative of the diversity of this taxonomic group [43]. Most amplified environmental sequences indeed do not tightly cluster with sequences retrieved from public databases (GenBank, CAZy, JGI, Broad Institute databases) (Figs. 3–4 & S2–S3). Furthermore, only the AA2 and GH5-5 sequences originating from Ascomycota and Basidiomycota form well separated clades allowing unambiguous assignation of homologous environmental sequences to one of these two broad fungal groups (Figs. 3 & S2). Absence of tight association between environmental sequences and reference ones was particularly pronounced for the AA2 gene family (Fig. 3). One likely explanation is that AA2 peroxidases have essentially been studied in the context of wood degradation which has promoted the genome sequencing of many wood degrading saprotrophic species [37, 40] and not of unrelated soil saprotrophs. Indeed, as recently evaluated, the PeroxiBase and GenBank databases contained 311 entries of class II (AA2) peroxidases from wood-decay fungi versus only 11 from litter-decomposing species [27]. These figures plead for increasing the sequencing effort of reference sequences from soil inhabiting fungi, especially those belonging to the Basidiomycota.

Finally, the results we obtained for the spruce and beech forest soil samples are coherent with the fungal taxonomic survey performed by Buée et al. [69] on the same forest plots. Indeed, Buée et al. [69] demonstrated a higher fungal diversity in the spruce compared to the beech soils (983 versus 581 operational taxonomic units), and latter, Buée et al. [56] also showed that saprotrophic macromycete species were more abundant under spruce than under beech where ectomycorrhizal taxa dominated. In the present study, we also identified higher numbers of expressed gene sequence clusters for all three studied gene families in the spruce samples (Table 2) which indeed may reflect a predominance and greater abundance of saprotrophic species in the corresponding forest soils.

Conclusions

In the present study, we designed primers targeting three functional fungal gene families encoding key enzymes involved in plant organic matter degradation (i.e. GH5-5, GH11 and AA2). We demonstrated their suitability for high-throughput sequencing using the Illumina MiSeq approach. We also evaluated two important

parameters associated with the high-throughput sequencing (i) at which percentage of identity should be clustered coding nucleotide sequences and (ii) the biological significance of singletons. Such an approach constitutes a robust method, which allows a detailed characterization of the diversity of soil expressed fungal genes involved in plant organic matter degradation and may lead to the discovery of patterns in gene expression by soil fungal communities that may go unnoticed using other traditional approaches.

Supporting Information

S1 Figure. Relation between the clustering threshold and the number of fungal endo-β-1,4-xylanase (GH11), cellulase (GH7) and Basidiomycota class II peroxidase (AA2) sequence clusters and of their "delta values". Evolution of the number of GH11 (A), GH7 (B) and AA2 (C) sequence clusters (N) expressed in soils (gray curve) and of their "delta values" (black curve) according to the percentage of dissimilarity used as cutoff for sequence clustering. "Delta values" represent the number of clusters at a cutoff of n% minus the values at n−1% ($\Delta = N_{n\%} - N_{n-1\%}$).
doi:10.1371/journal.pone.0116264.s001 (TIF)

S2 Figure. Phylogenetic positions of fungal endo-β-1,4-glucanase (GH5-5) amino-acid sequences. Fungal GH5-5 amino-acid sequences were deduced from the nucleotide sequences amplified from beech (BB) or spruce (BS) soil cDNAs. The Maximum-likelihood phylogenetic tree include all Sanger sequences amplified from the 2007 soil cDNA samples (BS2007 and BB2007) and all non-singleton sequence clusters detected by Illumina MiSeq sequencing of the 2010 soil cDNA samples (BS2010 and BB2010). Representative Ascomycota and Basidiomycota sequences are marked in red and blue, respectively, whereas the environmental sequences appear in green. Stars identify reference sequences obtained in the present study. Robustness of the tree topology was tested by bootstrap analysis (1000 replicates) and only bootstrap values ≥80 are given.
doi:10.1371/journal.pone.0116264.s002 (TIF)

S3 Figure. Phylogenetic positions of fungal cellulase (GH7) amino-acid sequences. Fungal GH7 amino-acid sequences were deduced from the nucleotide sequences amplified from beech (BB) or spruce (BS) soil cDNAs. The Maximum-likelihood phylogenetic tree include all Sanger sequences amplified from the 2007 soil cDNA samples (BS2007 and BB2007) and all non-singleton sequence clusters detected by Illumina MiSeq sequencing of the 2010 soil cDNA samples (BS2010 and BB2010). Representative Ascomycota and Basidiomycota sequences are marked in red and blue, respectively, whereas the environmental sequences appear in green. Stars identify reference sequences obtained in the present study. Robustness of the tree topology was tested by bootstrap analysis (1000 replicates) and only bootstrap values ≥80 are given.
doi:10.1371/journal.pone.0116264.s003 (TIF)

S1 Table. Reference sequences used for blastp search in NCBI and selected GenBank sequences for degenerate primer design.
doi:10.1371/journal.pone.0116264.s004 (DOCX)

S2 Table. Fungal species used to evaluate the degenerate primer efficiency and summary of sequencing results.
doi:10.1371/journal.pone.0116264.s005 (XLSX)

S3 Table. Analysis of the GH5-5, GH7, GH11 and AA2 sequences (Sanger sequencing) amplified from the 2007 forest soil cDNAs (Breuil Spruce (BS2007) and Breuil Beech (BB2007)).
doi:10.1371/journal.pone.0116264.s006 (DOCX)

Acknowledgments

We would like to thank both site managers (Jacques Ranger & Arnaud Legout) for fieldwork permits and Marc Buée (INRA Nancy) for soil sampling. We are indebted to colleagues who provided us fungal strains or fungal DNA (Stefania Daghino, Mariangela Girlanda, Harald Kellner, Erika Kothe, Ursula Kües, Francis Martin, Silvia Perotto, Philippe Silar, Christian Steinberg and Daniel Wipf). Many thanks to Harald Kellner and Guillaume Minard for valuable discussions.

Author Contributions

Conceived and designed the experiments: FB PL RM. Performed the experiments: FB CB EP LV PL. Analyzed the data: FB AD CB PL. Contributed reagents/materials/analysis tools: EP AD LV LFT. Contributed to the writing of the manuscript: FB CB LV EP AD LFT RM PL.

References

1. **Gifford RM** (2003) Plant respiration in productivity models: conceptualization, representation and issues for global terrestrial carbon-cycle research. Funct Plant Biol 30: 171–186.
2. **Lützow MV, Kögel-Knabner I, Ekschmitt K, Matzner E, Guggenberger G, et al.** (2006) Stabilization of organic matter in temperate soils: mechanisms and their relevance under different soil conditions – a review. Eur J Soil Sci 57: 426–445.
3. **Šnajdr J, Valášková V, Merhautová V, Herinková J, Cajthaml T, et al.** (2008) Spatial variability of enzyme activities and microbial biomass in the upper layers of *Quercus petraea* forest soil. Soil Biol Biochem 40: 2068–2075.
4. **Hättenschwiler S, Tiunov AV, Scheu S** (2005) Biodiversity and litter decomposition in terrestrial ecosystems. Annu Rev Ecol Evol Syst 36: 191–218.
5. **Sinsabaugh RL, Carreiro MM, Alvarez S** (2002) Enzyme and microbial dynamics of litter decomposition. In: Burns RG, Dick RP, editors. Enzymes in the environment. Marcel Deeker Inc., New York, USA, pp. 249–265.
6. **Sarkar P, Bosneaga E, Auer M** (2009) Plant cell walls throughout evolution: towards a molecular understanding of their design principles. J Exp Bot 13: 3615–3635.
7. **Austin AT, Ballaré CL** (2009) Dual role of lignin in plant litter decomposition in terrestrial ecosystems. Proc Natl Acad Sci USA 107: 4618–4622.

8. **de Boer W, Folman LB, Summerbell RC, Boddy L** (2005) Living in a fungal world: impact of fungi on soil bacterial niche development. FEMS Microbiol Rev 29: 795–811.

9. **Valášková V, Šnajdr J, Bittner B, Cajthaml T, Merhautová V, et al.** (2007) Production of lignocelluloses-degrading enzymes and degradation of leaf litter by saprotrophic basidiomycetes isolated from *Quercus petraea* forest. Soil Biol Biochem 39: 2651–2660.

10. **Baldrian P, Voříšková J, Dobiášová P, Merhautová V, Lisá L, et al.** (2011) Production of extracellular enzymes and degradation of biopolymers by saprotrophic microfungi from the upper layers of forest soil. Plant Soil 338: 111–125.

11. **Schneider T, Keiblinger KM, Schmid E, Sterflinger-Gleixner K, Ellersdorfer G, et al.** (2012) Who is who in litter decomposition? Metaproteomics reveals major microbial players and their biogeochemical functions. ISME J 6: 1749–1762.

12. **Zhou D, Hyde K** (2001) Host-specificity, host-exclusivity, and host-recurrence in saprobic fungi. Mycol Res 105: 1449–1457.

13. **Tedersoo L, Jairus T, Horton BM, Abarenkov K, Suvi T, et al.** (2008) Strong host preference of ectomycorrhizal fungi in a Tasmanian wet sclerophyll forest as revealed by DNA barcoding and taxon-specific primers. New Phytol 180: 479–490.

14. **Richard F, Roy M, Shahin O, Sthultz C, Duchemin M, et al.** (2011) Ectomycorrhizal communities in a Mediterranean forest ecosystem dominated by *Quercus ilex*: seasonal dynamics and response to drought in the surface organic horizon. Ann For Sci 68: 57–68.

15. **Šnajdr J, Cajthaml T, Valášková V, Merhautová V, Petránková M, et al.** (2011) Transformation of *Quercus petraea* litter: successive changes in litter chemistry are reflected in differential enzyme activity and changes in the microbial community composition. FEMS Microbiol Ecol 75: 291–303.

16. **Trap J, Laval K, Akpa-Vinceslas M, Gangneux C, Bureau F, et al.** (2011) Humus macro-morphology and soil microbial community changes along a 130-yr-old *Fagus sylvatica* chronosequence. Soil Biol Biochem 43: 1553–1562.

17. **Zhang D, Hui D, Luo Y, Zhou G** (2008) Rates of litter decomposition in terrestrial ecosystems: global patterns and controlling factors. J Plant Ecol 1: 85–93.

18. **Kellner H, Zak DR, Vandenbol M** (2010) Fungi unearthed: Transcripts encoding lignocellulolytic and chitinolytic enzymes in forest soil. PLoS One 5: e10971.

19. **Hao JJ, Tian XJ, Song FQ, He XB, Zhang ZJ, et al.** (2006) Involvement of lignocellulolytic enzymes in the decomposition of leaf litter in a subtropical forest. J Eukaryot Microbiol 53: 193–198.

20. **Osono T, Takeda H** (2006) Fungal decomposition of *Abies* needle and *Betula* leaf litter. Mycologia 98: 172–179.

21. **Herculano PN, Lima DMM, Fernandes MJS, Neves RP, Souza-Motta CM, et al.** (2011) Isolation of cellulolytic fungi from waste of Castor (*Ricinus communis* L.). Curr Microbiol 62: 1416–1422.

22. **Luis P, Kellner H, Zimdars B, Langer U, Martin F, et al.** (2005) Patchiness and spatial distribution of lacccase genes of ectomycorrhizal, saprotrophic, and unknown basidiomycetes in the upper horizons of a mixed forest cambisol. Microb Ecol 50: 570–579.

23. **Edwards IP, Upchurch RA, Zack DR** (2008) Isolation of fungal cellobiohydrolase I genes from sporocarps and forest soils by PCR. Appl Environ Microbiol 74: 3481–3489.

24. **Bödeker ITM, Nygren CMR, Taylor AFS, Olson A, Lindahl BD** (2009) ClassII peroxidase-encoding genes are present in a phylogenetically wide range of ectomycorrhizal fungi. ISME J 12: 1387–1395.

25. **Kellner H, Luis P, Schlitt B, Buscot F** (2009) Temporal changes in diversity and expression patterns of fungal laccase genes within the organic horizon of a brown forest soil. Soil Biol Biochem 41: 1380–1389.

26. **Bödeker ITM, Clemmensen KE, de Boer W, Martin F, Olson A, et al.** (2014) Ectomycorrhizal *Cortinarius* species participate in enzymatic oxidation of humus in northern forest ecosystems. New Phytol 203: 245–256.

27. **Kellner H, Luis P, Pecyna MJ, Barbi F, Kapturska D, et al.** (2014) Widespread occurrence of expressed fungal secretory peroxidases in forest soils. PLoS One 9: e95557.

28. **Baldrian P** (2006) Fungal laccases - occurrence and properties. FEMS Microbiol Rev 30: 215–242.

29. **Luis P, Walther G, Kellner H, Martin F, Buscot F** (2004) Diversity of laccase genes from basidiomycetes in forest soil. Soil Biol Biochem 36: 1025–1036.

30. **Edwards IP, Zak DR, Kellner H, Eisenlord SD, Pregitzer KS** (2011) Simulated atmospheric N deposition alters fungal community composition and suppresses ligninolytic gene expression in a northern hardwood forest. PLoS One 6: e20421.

31. **Weber CF, Zak DR, Hungate BA, Jackson RB, Vilgalys R, et al.** (2011) Responses of soil cellulolytic fungal communities to elevated atmospheric CO_2 are complex and variable across five ecosystems. Environ Microbiol 13: 2778–2793.

32. **Baldrian P, Kolařík M, Štursová M, Kopecký J, Valášková V, et al.** (2012) Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. ISME J 6: 248–258.

33. **Weber CF, Balasch MM, Gossage Z, Porras-Alfaro A, Kuske CR** (2012) Soil fungal cellobiohydrolase I gene (*cbhI*) composition and expression in a loblolly pine plantation under conditions of elevated atmospheric CO_2 and nitrogen fertilization. Appl Environ Microbiol 78: 3950–3957.

34. **Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B** (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42: 490–495.

35. **Asperborg H, Coutinho PM, Wang Y, Brumer H III, Henrissat B** (2012) Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). BMC Evol Biol 12: 186.

36. **Yoon JJ, Cha CJ, Kim YS, Kim W** (2008) Degradation of cellulose by the major endoglucanase produced from the brown-rot fungus *Fomitopsis pinicola*. Biotechnol Lett 30: 1373–1378.

37. **Martinez D, Challacombe J, Morgenstern I, Hibbett D, Schmoll M, et al.** (2009) Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. Proc Natl Acad Sci USA 106: 1954–1959.

38. **MacDonald J, Doering M, Canam T, Gong Y, Guttman DS, et al.** (2011) Transcriptomic responses of the softwood-degrading white-rot fungus *Phanerochaete carnosa* during growth on coniferous and deciduous Wood. Appl Environ Microbiol 77: 3211–3218.

39. **Schädel C, Blöchl A, Richter A, Hoch G** (2010) Quantification of monosaccharide composition of hemicelluloses from different plant functional types. Plant Physiol Bioch 48: 1–8.

40. **Floudas D, Binder M, Riley R, Barry K, Blanchette RA, et al.** (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336: 1715–1719.

41. **Sinsabaugh RL** (2010) Phenol oxidase, peroxidase and organic matter dynamics of soil. Soil Biol Biochem 42: 391–404.

42. **Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL** (2008) GenBank. Nucleic Acids Res 36: 25–30.

43. **Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, et al.** (2014) MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Res 42: 699–704.

44. **Möller EM, Bahnweg G, Sandermann H, Geiger HH** (1992) A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. Nucleic Acids Res 20: 6115–6116.

45. **Ranger J, Andreux BSF, Berthelin BP, Boudot JP, Bréchet C, et al.** (2004) Effet des substitutions d'essence sur le fonctionnement organo-minéral de l'écosystème forestier, sur les communautés microbiennes et sur la diversité des communautés fongiques mycorhiziennes et saprophytes (cas du dispositif expérimental de Breuil - Morvan). Final report of contract INRA-GIP Ecofor 2001–24. Champenoux, No. INRA 1502A. INRA BEF Nancy, France.

46. **Damon C, Barroso G, Férandon C, Ranger R, Fraissinet-Tachet L, et al.** (2010) Performance of the *COX1* gene as a marker for the study of metabolically active Pezizomycotina and Agaricomycetes fungal communities from the analysis of soil RNA. FEMS Microbiol Ecol 74: 693–705.

47. **Masella A, Bartram A, Truszkowski J, Brown D, Neufeld J** (2012) PANDAseq: paired-end assembler for Illumina sequences. BMC Bioinformatics 13: 31.

48. **Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, et al.** (2009) Introducing mothur: opensource, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75: 7537–7541.

49. **Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R** (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27: 2194–2200.

50. **Edgar RC** (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32: 1792–1797.

51. **Fawal N, Li Q, Savelli B, Brette M, Passaia G, et al.** (2013) PeroxiBase: a database for large-scale evolutionary analysis of peroxidases. Nucleic Acids Res 41: 441–444.

52. **Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, et al.** (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59: 307–21.

53. **Whelan S, Goldman N** (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18: 691–699.

54. **Gouy M, Guindon S, Gascuel O** (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27: 221–224.

55. **Voříšková J, Baldrian P** (2013) Fungal community on decomposing leaf litter undergoes rapid successional changes. ISME J 7: 477–486.

56. **Buée M, Maurice JP, Zeller B, Andrianarisoa S, Ranger J, et al.** (2011) Influence of tree species on richness and diversity of epigeous fungal communities in a French temperate forest stand. Fungal Ecol 4: 22–31.

57. **Manzoni S, Schimel JP, Porporato A** (2012) Responses of soil microbial communities to water stress: results from a meta-analysis. Ecology 93: 930–938.

58. **Hadziavdic K, Lekang K, Lanzen A, Jonassen I, Thompson EM, et al.** (2014) Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. PLoS One 9: e87624.

59. **Levasseur A, Saloheimo M, Navarro D, Andberg M, Pontarotti P, et al.** (2010) Exploring laccase-like multicopper oxidase genes from the ascomycete *Trichoderma reesei*: a functional, phylogenetic and evolutionary study. BMC Biochem 11: 32.

60. **Palomares-Rius JE, Hirooka Y, Tsai IJ, Masuya H, Hino A, et al.** (2014) Distribution and evolution of glycoside hydrolase family 45 cellulases in nematodes and fungi. BMC Evol Biol 14: 69.

61. **Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, et al.** (2012) Performance comparison of benchtop high-throughput sequencing platforms. Nat Biotechnol 30: 434–562.

62. **Urich T, Lanzen A, Qi J, Huson DH, Schleper C, et al.** (2008) Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. PLoS One 3: e2527.

63. **Bailly J, Fraissinet-Tachet L, Verner MC, Debaud JC, Lemaire M, et al.** (2007) Soil eukaryotic functional diversity, a metatranscriptomic approach. ISME J 1: 632–642.

64. **Damon C, Lehembre F, Oger-Desfeux C, Luis P, Ranger J, et al.** (2012) Metratranscriptomics reveals the diversity of genes expressed by eukaryotes in forest soils. PLoS One 7: e28967.

65. **Wall DP, Hirsh AE, Fraser HB, Kumm J, Giaever G, et al.** (2005) Functional genomic analysis of the rates of protein evolution. Proc Natl Acad Sci USA 102: 5483–5488.

66. **Wolfe BE, Tulloss RE, Pringle A** (2012) The irreversible loss of a decomposition pathway marks the single origin of an ectomycorrhizal symbiosis. PLoS One 7: e39597.

67. **Schmidt PA, Bálint M, Greshake B, Bandow C, Römbke J, et al.** (2013) Illumina metabarcoding of a soil fungal community. Soil Biol Biochem 65: 128–132.

68. **Wubet T, Christ S, Schöning I, Boch S, Gawlich M, et al.** (2012) Differences in soil fungal communities between european beech (*Fagus sylvatica* L.) dominated forests are related to soil and understory vegetation. PLoS One 7: e47500.

69. **Buée M, Reich M, Murat C, Morin E, Nilsson RH, et al.** (2009) 454 pyrosequencing analyses of forest soils reveal an unexpectedly high fungal diversity. New Phytol 184: 449–456.

70. **Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD** (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol 79: 5112–5120.

Supporting information

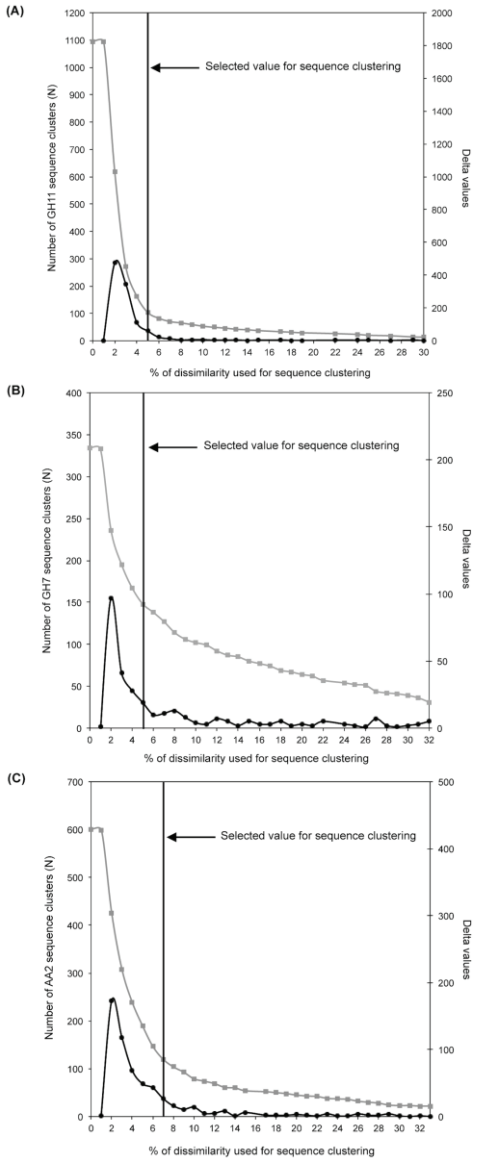

Figure S1. Relation between the clustering threshold and the number of fungal endo-β-1,4-xylanase (GH11), cellulase (GH7) and Basidiomycota class II peroxidase (AA2) sequence clusters and of their "delta values". Evolution of the number of GH11 (A), GH7 (B) and AA2 (C) sequence clusters (N) expressed in soils (gray curve) and of their "delta values" (black curve) according to the percentage of dissimilarity used as cutoff for sequence clustering. "Delta values" represent the number of clusters at a cutoff of n% minus the values at n−1% ($\Delta = N_{n\%} - N_{n-1\%}$).

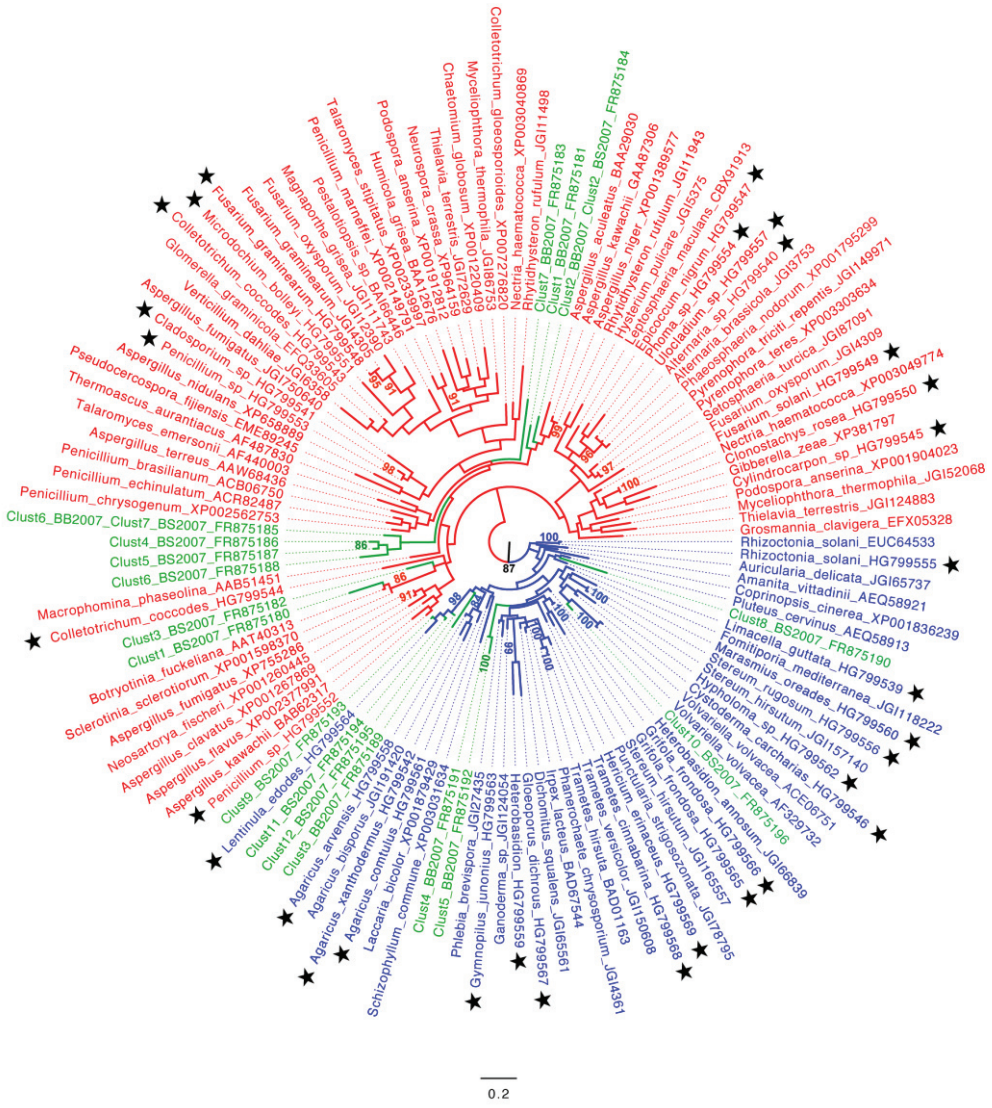

Figure S2. Phylogenetic positions of fungal endo-β-1,4-glucanase (GH5-5) amino-acid sequences. Fungal GH5-5 amino-acid sequences were deduced from the nucleotide sequences amplified from beech (BB) or spruce (BS) soil cDNAs. The Maximum-likelihood phylogenetic tree include all Sanger sequences amplified from the 2007 soil cDNA samples (BS2007 and BB2007) and all non-singleton sequence clusters detected by Illumina MiSeq sequencing of the 2010 soil cDNA samples (BS2010 and BB2010). Representative Ascomycota and Basidiomycota sequences are marked in red and blue, respectively, whereas the environmental sequences appear in green. Stars identify reference sequences obtained in the present study. Robustness of the tree topology was tested by bootstrap analysis (1000 replicates) and only bootstrap values ≥80 are given.

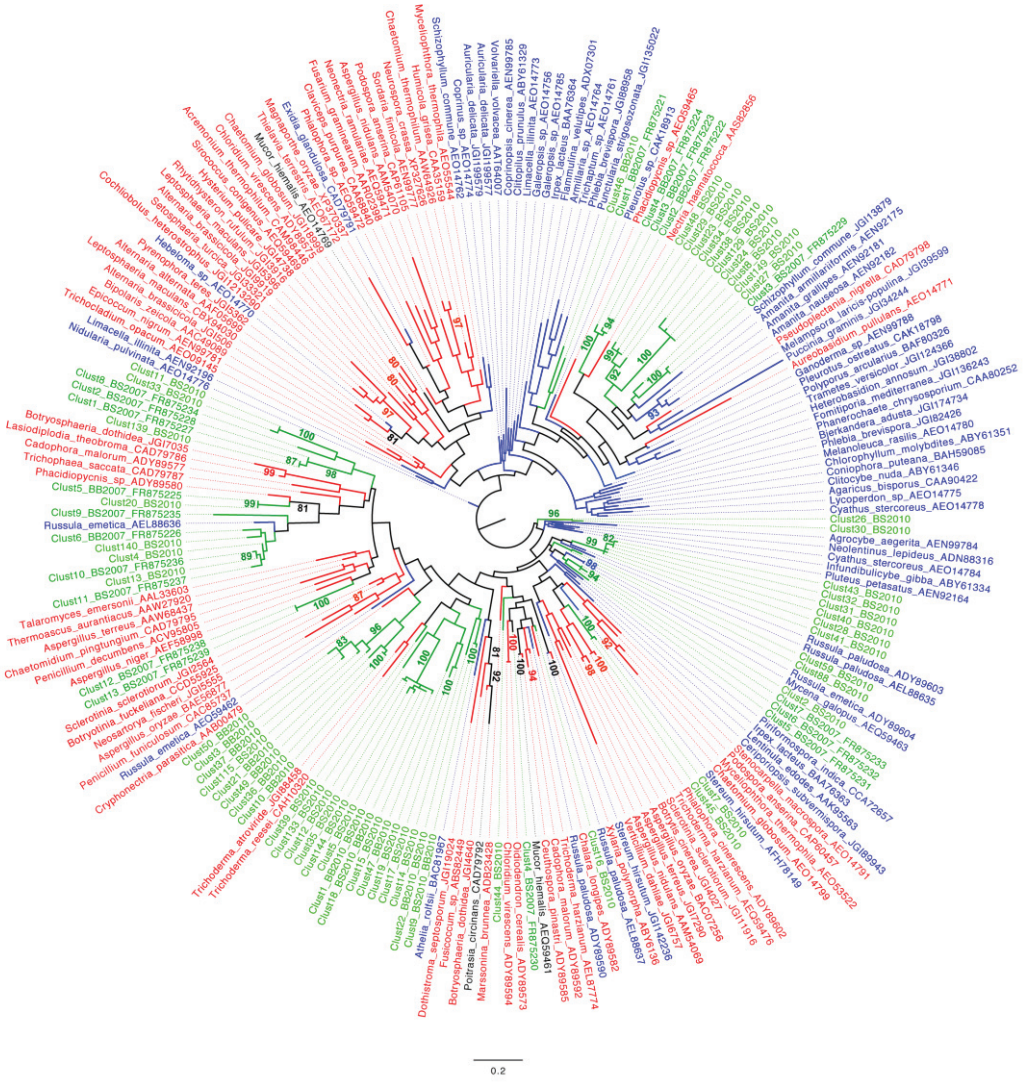

Figure S3. Phylogenetic positions of fungal cellulase (GH7) amino-acid sequences. Fungal GH7 amino-acid sequences were deduced from the nucleotide sequences amplified from beech (BB) or spruce (BS) soil cDNAs. The Maximum-likelihood phylogenetic tree include all Sanger sequences amplified from the 2007 soil cDNA samples (BS2007 and BB2007) and all non-singleton sequence clusters detected by Illumina MiSeq sequencing of the 2010 soil cDNA samples (BS2010 and BB2010). Representative Ascomycota and Basidiomycota sequences are marked in red and blue, respectively, whereas the environmental sequences appear in green. Stars identify reference sequences obtained in the present study. Robustness of the tree topology was tested by bootstrap analysis (1000 replicates) and only bootstrap values ≥80 are given.

Table S1. Reference sequences used for blastp search in NCBI and selected GenBank sequences for degenerate primer design.

Enzyme families/subfamilies	Accession numbers	Fungal groups	Fungal species	Conserved protein regions 1	Conserved protein regions 2	Conserved protein sequences selected for primer design
GH5-5	BAD01163 [1]	Basidiomycota	*Trametes hirsuta*	QMHQYLD	WWAAGPW	(Q/E)MHQYLD
	XP_008039932	Basidiomycota	*Trametes versicolor*	QMHQYLD	WWAAGPW	&
	BAD67544	Basidiomycota	*Irpex lacteus*	EMHQYLD	WWAAGPW	WWAAGPW
	ETW82970	Basidiomycota	*Heterobasidion irregulare*	EMHQYLD	WWAAGPW	
	XP_007368133	Basidiomycota	*Dichomitus squalens*	QMHQYLD	WWAAGPW	
	KDR76291	Basidiomycota	*Galerina marginata*	QMHQYLD	WWAAGPW	
	XP_007378214	Basidiomycota	*Punctularia strigosozonata*	EMHQYLD	WWAAGQW	
	XP_007393231	Basidiomycota	*Phanerochaete carnosa*	EMHQYLD	WWAAGPW	
	XP_007347849	Basidiomycota	*Auricularia delicata*	EMHQYLD	WWAAGPW	
	XP_007300340	Basidiomycota	*Stereum hirsutum*	EMHQYLD	WWAAGQW	
	KDQ10377	Basidiomycota	*Botryobasidium botryosum*	EMHQYLD	WWAAGPW	
	CDJ79817	Basidiomycota	*Leucoagaricus gongylophorus*	EMHQYLD	WWAAGPW	
	XP_007263372	Basidiomycota	*Fomitiporia mediterranea*	EMHQYLD	WWAAGPW	
	KEP51038	Basidiomycota	*Rhizoctonia solani*	EMHQYLD	WWAAGPW	
	XP_001879429	Basidiomycota	*Laccaria bicolor*	EMHQYLD	WWAAGPW	
	XP_006459448	Basidiomycota	*Agaricus bisporus*	QMHQYLD	WWSAGPW	
	CCX05395	Ascomycota	*Pyronema omphalodes*	EMHQYLD	WWAAGPW	
	XP_003031634	Basidiomycota	*Schizophyllum commune*	EMHQYLD	WWAAGPW	
	XP_007304563	Basidiomycota	*Stereum hirsutum*	EMHQYLD	WWAAGPW	
	XP_007851566	Basidiomycota	*Moniliophthora roreri*	EMHQYLD	WWAAGPW	
	KDQ27083	Basidiomycota	*Pleurotus ostreatus*	EMHQYLD	WWAAGPW	
	XP_002840457	Ascomycota	*Tuber melanosporum*	EMHQYLD	WWAAGPW	
	XP_001836239	Basidiomycota	*Coprinopsis cinerea*	QMHQYLD	WWAAGPW	
	XP_003654855	Ascomycota	*Thielavia terrestris*	EMHQYLD	WWGAGPW	
	XP_007581661	Ascomycota	*Neofusicoccum parvum*	QMHQYLD	WWAAGPW	
	EGX49988	Ascomycota	*Arthrobotrys oligospora*	EMHQYLD	WWAAGPW	
	XP_001934618	Ascomycota	*Pyrenophora tritici-repentis*	QMHQYLD	WWAAGPW	

	EFX05328	Ascomycota	*Grosmannia clavigera*	EMHQYLD	WWGAGPW	
	XP_003049774	Ascomycota	*Nectria haematococca*	EMHQYLD	WWAAGPW	
	XP_003663700	Ascomycota	*Myceliophthora thermophila*	EMHQYLD	WWAAGPW	
	XP_007290718	Ascomycota	*Marssonina brunnea*	EMHQYLD	WWAGGPW	
	AAL33630	Ascomycota	*Rasamsonia emersonii*	EMHQYLD	WWAAGPW	
	XP_381797	Ascomycota	*Fusarium graminearum*	QMHQYLD	WWAAGPW	
	EPS30243	Ascomycota	*Penicillium oxalicum*	EMHQYLD	WWAAGPW	
	EFQ31119	Ascomycota	*Colletotrichum graminicola*	EMHQYLD	WWAAGPW	
	AHY00945	Ascomycota	*Aspergillus niger*	EMHQYLD	WWAAGPW	
	EME41652	Ascomycota	*Dothistroma septosporum*	QMHQYLD	WWAAGPW	
	ESZ99399	Ascomycota	*Sclerotinia borealis*	EFHQYLD	WWAAGPW	
	AFY98622	Ascomycota	*Bispora antennata*	EMHQYLD	WWGGGPW	
	AAT40313	Ascomycota	*Botrytis cinerea*	EFHQYLD	WWGAGPW	
	XP_001904023	Ascomycota	*Podospora anserina*	QMHQYLD	WWGGGPW	
	XP_002149791	Ascomycota	*Talaromyces marneffei*	EMHQYLD	FWAAGPW	
	EKG11044	Ascomycota	*Macrophomina phaseolina*	QMHQYLD	WWAAGPW	
	KEQ63578	Ascomycota	*Aureobasidium melanogenum*	QMHQYLD	WWAAGPW	
	AEQ58918	Basidiomycota	*Amanita prairiicola*	QMHQYLD	WWAAGPW	
	AEQ58912	Basidiomycota	*Limacella illinita*	QMHQYLD	WWAAGPW	
	AEQ58913	Basidiomycota	*Pluteus cervinus*	EMHQYLD	WWAAGPW	
	AEQ58915	Basidiomycota	*Pluteus petasatus*	EMHQYLD	WWAAGPW	
	AEQ58914	Basidiomycota	*Volvariella volvacea*	EMHQYLD	WWAAGPW	
	XP_007795848	Ascomycota	*Eutypa lata*	DIHEYLD	WWAAGPW	
	XP_006692513	Ascomycota	*Chaetomium thermophilum*	EMHQYLD	WWAAGPW	
	KDQ30980	Basidiomycota	*Pleurotus ostreatus*	DVHKYLD	VWSAGSF	
	XP_008084003	Ascomycota	*Glarea lozoyensis*	DIHEYLD	AWAAGPF	
GH11	AAL04152 [1]	Basidiomycota	*Lentinula edodes*	GKGWNPG	QYWSIRQ	GKGWNPG
	XP_007397319	Basidiomycota	*Phanerochaete carnosa*	GKGWNPG	QYWSIRS	&
	ABZ88798	Basidiomycota	*Phanerochaete chrysosporium*	GKGWNPG	QYWSIRS	QYWS(I/V)RQ
	ADZ99359	Ascomycota	*Phialophora* sp.	GKGWNPG	QYWSVRR	
	XP_007379856	Basidiomycota	*Punctularia strigosozonata*	GKGWNPG	QYWSIRS	

XP_007300161	Basidiomycota	*Stereum hirsutum*	GKGWNPG	QYWSVRQ
XP_007343831	Basidiomycota	*Auricularia delicata*	GKGWNPG	QYFSIRQ
XP_001258363	Ascomycota	*Neosartorya fischeri*	GKGWNPG	QYWSVRT
EME39847	Ascomycota	*Dothistroma septosporum*	GKGWNPG	QYWSIRT
CAC15487	Ascomycota	*Talaromyces funiculosus*	GKGWNPG	QYWSVRT
XP_001389848	Ascomycota	*Aspergillus niger*	GKGWNPG	QYWSVRT
XP_008078781	Ascomycota	*Glarea lozoyensis*	GKGWNPG	QYWSVRQ
BAE71133	Ascomycota	*Penicillium citrinum*	GKGWNPG	QYWSVRQ
EMD39588	Basidiomycota	*Ceriporiopsis subvermispora*	GKGWNPG	QYWSIRS
EFQ30380	Ascomycota	*Colletotrichum graminicola*	GKGWNPG	QYWSVRQ
EMF09157	Ascomycota	*Sphaerulina musiva*	GKGWNPG	QYWSIRT
ADF27784	Ascomycota	*Morchella spongiola*	GKGWSPG	QYWSVRK
CCA76116	Basidiomycota	*Piriformospora indica*	GKGWNPG	QFWSVRQ
BAO51921	Ascomycota	*Talaromyces cellulolyticus*	GKGWNPG	QYWSVRT
KDR71791	Basidiomycota	*Galerina marginata*	GKGWNPG	QYWSVRQ
EJT75607	Ascomycota	*Gaeumannomyces graminis*	GKGWNPG	QYWAIRT
ETI21992	Ascomycota	*Cladophialophora carrionii*	GKGWNPG	QYWSIRT
XP_003050975	Ascomycota	*Nectria haematococca*	GKGWNPG	QYWSVRR
EGU83183	Ascomycota	*Fusarium oxysporum*	GKGWMPG	QYWSVRR
AAP83925	Ascomycota	*Trichoderma viride*	GKGWNPG	QYWSVRR
XP_008028976	Ascomycota	*Setosphaeria turcica*	GKGWNPG	QFWSVRT
164562254	Basidiomycota	*Pleurotus ostreatus*	GKGWNPG	QYWSVRT
ESZ90340	Ascomycota	*Sclerotinia borealis*	GKGWAVG	QYWSVRT
ABE02800	Ascomycota	*Verticillium dahliae*	GKGWNPG	QYWSVRT
XP_001910545	Ascomycota	*Podospora anserina*	GKGWNPG	QFWSVRR
XP_003662402	Ascomycota	*Myceliophthora thermophila*	GKGWNPG	QFWSVRT
ABG33753	Ascomycota	*Alternaria* sp.	GKGWNPG	QYWSVRT
XP_007912948	Ascomycota	*Togninia minima*	GKGWNPG	QYWSIRQ
XP_003837356	Ascomycota	*Leptosphaeria maculans*	GKGWNPG	QYWSVRQ
XP_007292301	Ascomycota	*Marssonina brunnea*	GKGYKPG	QYWSVRR
KEQ91156	Ascomycota	*Aureobasidium subglaciale*	GKGWNPG	QYWSVRQ
AHC72381	Ascomycota	*Humicola insolens*	GKGWNPG	QYWSVRR
AAZ03776	Ascomycota	*Botrytis cinerea*	GKGWAVG	QYWSVRT

	XP_001941158	Ascomycota	*Pyrenophora tritici-repentis*	GKGWNPG	QYWSVRT	
	XP_006690651	Ascomycota	*Chaetomium thermophilum*	GKGWNPG	QYWSVRR	
	EHK25898	Ascomycota	*Trichoderma virens*	GKGWNPG	QYWSVRR	
	GAD94166	Ascomycota	*Byssochlamys spectabilis*	GKGWSTG	QYWSVRQ	
	ABG37634	Ascomycota	*Acrophialophora nainiana*	GKGWNPG	QYWSVRQ	
	CAA76570	Ascomycota	*Claviceps purpurea*	GRGWNPG	QYWSIRR	
	EUN22843	Ascomycota	*Bipolaris victoriae*	GKGRNPG	QYWSVRQ	
	XP_007857498	Ascomycota	*Moniliophthora roreri*	GKGWNPG	QFWSVRN	
	XP_003715734	Ascomycota	*Magnaporthe oryzae*	GKGWNPG	QFWSVRR	
AA2	AAA34049 [1]	Basidiomycota	*Trametes versicolor*	GGGADGS	PFDSTP	GGGADGS & PFDSTP
	XP_008043740	Basidiomycota	*Trametes versicolor*	GGGADGS	PFDSTP	
	XP_008043625	Basidiomycota	*Trametes versicolor*	GGGADGS	PFDSTP	
	AEJ37998	Basidiomycota	*Polyporus brumalis*	GGGADGS	PFDSTP	
	AEX01147	Basidiomycota	*Lenzites gibbosa*	GGGADGS	PFDSTP	
	BAE79812	Basidiomycota	*Spongipellis* sp.	GGGADGS	PFDSTP	
	1906181A	Basidiomycota	*Bjerkandera adusta*	GGGADGS	PFDSTP	
	ADK26471	Basidiomycota	*Hericium erinaceus*	GGGADGS	PFDSTP	
	BAG85350	Basidiomycota	*Phanerochaete sordida*	GGGADGS	PFDSTP	
	XP_007365204	Basidiomycota	*Dichomitus squalens*	GGGADGS	PFDSTP	
	AGO86670	Basidiomycota	*Irpex lacteus*	GGGADGS	PFDSTP	
	BAG49629	Basidiomycota	*Ceriporiopsis* sp.	GGGADGS	PFDSTP	
	ETW80422	Basidiomycota	*Heterobasidion irregulare*	GGGADGS	PFDSTP	
	XP_007266673	Basidiomycota	*Fomitiporia mediterranea*	GGGADGS	PFDSTP	
	ACM47219	Basidiomycota	*Pleurotus ostreatus*	GGGADGS	PFDSTP	
	ETW82129	Basidiomycota	*Heterobasidion irregulare*	SGGADGS	PFDSTP	
	XP_007845338	Basidiomycota	*Moniliophthora roreri*	GGGADGS	PFDSTP	
	AFR44747	Basidiomycota	*Volvariella volvacea*	GGGADGS	PFDSTP	
	AFK91531	Basidiomycota	*Cerrena unicolor*	GGGADGS	PFDSTP	
	ADW41626	Basidiomycota	*Agrocybe praecox*	GGGADGS	PFDSTP	
	ADK60913	Basidiomycota	*Trametes cinnabarina*	GGGADGS	PFDSTP	
	ADK60911	Basidiomycota	*Trametes cinnabarina*	GGGADGS	PFDSTP	
	XP_006460927	Basidiomycota	*Agaricus bisporus*	GGGADGS	PFDSTP	
	XP_007309309	Basidiomycota	*Stereum hirsutum*	GGGADGS	PFDSTP	
	ACA48488	Basidiomycota	*Ganoderma lucidum*	GGGADGS	PFDSTP	

KDR70719	Basidiomycota	*Galerina marginata*	GGGADGS	PFDTSP
BAE46585	Basidiomycota	*Trametopsis cervina*	GGGADGS	PFDSTP
ABT17196	Basidiomycota	*Phanerochaete chrysosporium*	GGGADGS	PFDSTP
XP_007353069	Basidiomycota	*Auricularia delicata*	GGGADGS	PFDSTP
AAW59419	Basidiomycota	*Phlebia radiata*	GGGADGS	PFDSTP
XP_007382780	Basidiomycota	*Punctularia strigosozonata*	GGGADGS	PFDSTP

[1] CAZy reference sequences used for blastp search

Table S3. Analysis of the GH5-5, GH7, GH11 and AA2 sequences (Sanger sequencing) amplified from the 2007 forest soil cDNAs (Breuil Spruce (BS2007) and Breuil Beech (BB2007)).

	BS2007 GH5-5	**BB2007 GH5-5**	**BS2007 GH7**	**BB2007 GH7**	**BS2007 GH11**	**BB2007 GH11**	**BS2007 AA2**	**BB2007 AA2**
No. of sequenced clones	48	48	48	48	48	48	48	48
No. of bad quality sequences	9	14	14	5	11	4	7	11
No. of chimeric sequences	0	0	0	0	0	0	0	22
No. of fungal sequences	39	34	34	43	37	44	41	15
No. of cluster (singletons)[1]	12 (6)	7 (3)	13 (4)	6 (2)	16 (8)	12 (6)	17 (11)	6 (4)

[1] nucleotide sequences were clustered at 100% identity.

Chapter VI

Final remarks & future perspectives

The last decade has seen a growing number of studies testing how belowground microbial communities are distributed in space and time (e.g. Fierer & Jackson 2006, Bates *et al.* 2013, Talbot *et al*. 2014, Tedersoo *et al.* 2014), how they respond to global changes and what are the consequences of changes in soil biodiversity in term of plant community dynamics, aboveground trophic interactions and biogeochemical cycles (Bardgett & van der Putten 2014). Moreover soil biodiversity research is now entering a new era. Indeed, awareness is growing among scientists but also policy makers of the importance of soil biodiversity for the supply of ecosystems goods and services to human society. As such, the year 2015 has been declared by the United Nation as the International Year of Soils. Among soil microorganisms, fungi play pivotal roles in soil ecosystems functioning but also contribute to the direct well-being of mankind. In this context, it has been advocated that fungal products constitute essential building block for change towards a more sustainable future for our planet (Lange 2010).

New experimental approaches and tools are now available to interrogate the biology of soil fungal communities, their dynamics, their ecological functions and their roles in ecosystem functioning. Some of these approaches (metabarcoding, metatranscriptomics) and tools (high-throughput sequencing and associated data-handling softwares) have been implemented in this thesis which aimed at understanding the impact of different global changes on soil fungal community taxonomic and functional diversity.

From the results obtained, it was suggested that fungal responses to environmental changes may not be readily predictable, as they are rather context-dependent.

In *Chapter II*, a soil DNA metabarcoding approach allowed to characterise at a fine taxonomic level specific members of the arbuscular mycorrhizal fungal community in different agricultural soils across Europe. These organisms were found to respond in a site-dependent way to land use changes. However, to really appreciate the extent of different land uses on fungal communities, metabarcoding could be extended to soil-extracted RNA instead of DNA. Indeed, as reported by Baldrian *et al.* (2012), potential (DNA) and active (RNA) fungal communities may significantly differ. Working on RNA may more accurately reflect the diversity of active fungal species represented as metabolically active hyphae exploring the soil matrix, while soil DNA may include species only present as quiescent spores. Regarding *Chapter II*, another perspective could also be to extend the analysis to other fungal taxonomic groups (which in agricultural soils encompass "plant-independent" saprotrophic as well as pathogenic species) and compare their responses to the land-use factors to the responses recorded with the plant-dependent AM fungi. As in the case of *Chapter III*, the study presented in *Chapter II* indicated that a large part of the observed variation between communities could not

be explained by the different variables retained for the analyses. We suggest that the community analysis presented in both *Chapters II* and *III* would benefit from the analysis of additional soil samples collected all along the year, at regular (weekly, monthly) intervals in order to integrate in the analyses the impact of additional variables such as plant phenology and short-term changes in abiotic meteorological parameters (temperature, humidity...) which are known to be important drivers of microbial metabolism.

Plant organic matter decomposition is a complex environmental process. Its understanding, especially in plant litter and soil, was previously limited to the measurement of crude enzyme activities from environmental samples and/or of studies measurement of the decomposition rates of natural compounds. With the accumulation of gene sequences data for decomposition-related enzymes in public databases, it became possible in the late 2000s to design degenerate primers capable of amplifying the corresponding gene sequences from a wide range of fungi.

In *Chapter III* we implemented a parallel targeted-metatranscriptomics approach on three genes implicated in plant cell wall polymer degradation in order to appreciate the impact of a simulated climate change (partial rainfall removal) on the diversity of the soil saprotrophic communities. It must be stressed that although we could not detect any significant effect of rainfall reduction on the diversity of the studied genes (but for a single exception), we cannot exclude that this climate change does not have an impact which would affect other taxonomic or functional microbial groups or essential metabolic pathways. This is currently being tested in the Microbial Ecology laboratory in Lyon with the systematic shotgun sequencing of (i) the total soil RNA and (ii) the eukaryotic poly-A fraction from the different soil samples studied in *Chapter III*. As total soil RNA is dominated by ribosomal RNA (rRNA), its analysis will give indications on the global microbial (bacteria, archaea, eukarya) diversity of the soil samples while annotation of the poly-A mRNA sequence dataset (more than 10^9 sequences already available) will give an overview of all the different activities expressed by the eukaryotic community.

In *Chapter III* we were also confronted to the almost impossibility at assigning soil gene and protein sequences not only at the species level, but also at the genus, family, order and sometimes even higher taxonomic levels. This limitation will probably not be solved in the immediate future as it will require the systematic sequencing of homologous gene sequences from a large number of referenced species representatives of the soil ecosystem. In this respect, large scale genome sequencing programs, like the 1000 Fungal Genomes program (Grigoriev *et al.* 2014) will be of limited value as the selected taxa are chosen on a taxonomic basis, not an an ecological one.

We propose below a possible experimental strategy for the taxonomic assignation of the most abundant soil sequence clusters, which could represent the most active

(abundant) fungal saprotrophic species in the corresponding soils. This strategy is based on the observation and hypothesis that abundant clusters (i) are frequently found in different soil samples and (ii) have probably a higher probability to originate from species abundant in term of biomass than from low-abundance species. We therefore propose to quantify the relative abundance of the different fungal species in our soil samples by performing a high-throughput sequencing of the rDNA ITS region on the DNA extracted from our different soil samples. As opposed to enzyme coding sequences, the ITS has been approved as the universal barcode sequences for the fungi (Schoch *et al.* 2012) and soil sequences can often be assigned at the species- or genus-level. We then propose to collect/extract DNA from field-collected fruiting bodies or herbarium specimens or living cultures of the most abundant saprotrophic species thus identified, amplify by PCR and sequence their PCW degrading genes and match the sequences to sequences obtained from soil RNA. Although this strategy of "reverse ecology" could be cumbersome, it could allow us to identify saprotrophic keystone species contributing most to soil organic matter degradation in a specific forest ecosystem. Activities of these species could then be quantified by quantitative PCR using non-degenerated gene-specific primers. This strategy has already been initiated and we have obtained ITS MiSeq sequences from almost all soil samples studied in *Chapter III.*

More classic approaches for the analysis of decomposition process *in-situ* (enzyme assays on soil extracts, soil respiration) should also be implemented to correlate molecular data to ecosystem process. Climate change can indeed impact fungal decomposer community indirectly, by changing plant organic matter composition, and carbon allocation to the soil. Data on plant litter composition and carbon input to the soil should be integrated in our analyses. Moreover, as the activity of enzymes degrading plant polymers is extremely variable in responses to environmental variation, such as moisture, temperature...., as already mentioned above, frequent and regular soil samplings (e.g. at weekly intervals), followed by enzymatic/metagenomic/metatranscriptomic analyses should be perform to elucidate this extremely complex process in nature.

Finally, besides its power for the elucidation of natural processes, environmental genomics present an enormous potential for the discovery of novel bioactive compounds and biocatalysts from complex environmental matrices such as soil. Indeed, many of the most widely used antibiotics have come out from soil microorganisms and in a first metagenomic analysis, (Tringe *et al.* 2005) showed that antibiotic-coding genes were more abundant in a soil metagenome than in a sea water one. Penicillin came from *Penicillium*, a fungus found in soil, and vancomycin came from a soil actinomycete. Recently, researchers from the Northeastern University (USA) and NovoBiotic Pharmaceuticals and their colleagues have identified a new Gram-positive bacteria-targeting antibiotic from a soil sample

collected in Maine that can kill species including methicillin-resistant *Staphylococcus aureus* (MRSA) and *Mycobacterium tuberculosis*. Moreover, the researchers have not yet found any bacteria that are resistant to the antibiotic, called teixobactin (Ling *et al.* 2015).

The sequence capture technique, which was described in *Chapter IV*, has the potential to recover enzymes that may be employed in the biotech industry. Its application to other gene families could allow the enzymatic characterization of families of enzymes displaying more than one activity, such as for instance the Glycoside hydrolases 5 family, which includes sub-families with different known activities and others with no known activities (http://www.cazy.org/GH5.html). To do that, the protocol presented in *Chapter IV* can be rigorously followed but also modified by changing for instance the expression vector and host, which could be another yeast, a bacterium or even an animal cell.

References:

Baldrian, P., Kolařík, M., Stursová, M., Kopecký, J., Valášková, V., Větrovský, T., *et al.* (2012). Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. *ISME J.* 6:248–58.

Bardgett, R.D., van der Putten, W.H. (2014). Belowground biodiversity and ecosystem functioning. *Nature* 515:505–511.

Bates, S.T., Clemente, J.C., Flores, G.E., Walters, W.A., Parfrey, L.W., Knight, R., *et al.* (2013). Global biogeography of highly diverse protistan communities in soil. *Isme J.* 7:652–659.

Fierer, N., Jackson, R.B. (2006). The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci*. U. S. A. 103:626–631.

Grigoriev, IV., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R. *et al.* (2014). MycoCosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Res*. 42.

Lange, L. (2010). The importance of fungi for a more sustainable future on our planet. *Fungal Biol. Rev*. 24:90–92.

Ling, L.L., Schneider, T., Peoples, A.J., Spoering, A.L., Engels, I., Conlon, B.P., *et al.* (2015). A new antibiotic kills pathogens without detectable resistance. *Nature* 517:455–459.

Schoch, C. L.; Seifert, K. A.; Huhndorf, S.; *et al.* , Fungal Barcoding Consortium (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi *Proc. Natl. Acad. Sci. U. S. A.* 109(16): 6241-6246.

Talbot, J.M., Bruns, T.D., Taylor, J.W., Smith, D.P., Branco, S., Glassman, S.I., *et al.* (2014). Endemism and functional convergence across the North American soil mycobiome. *Proc. Natl. Acad. Sci. U. S. A.* 111:6341–6.

Tedersoo, L., Bahram, M., Polme, S., Koljalg, U., Yorou, N. S., Wijesundera, R., *et al.* (2014). Global diversity and geography of soil fungi. *Science* 346(6213):1256688–1256688.

Tringe, S.G, von Mering ,C., Kobayashi, A., Salamov, A.A., Chen, K., Chang, H.W., *et al.* (2005). Comparative metagenomics of microbial communities. *Science* 308:554–557.

Ringraziamenti/Acknowledgements/Remerciements

Che dire….questi tre anni son volati, mille persone ne hanno fatto parte..chi solo per un momento chi dall'inizio alla fine….

Tra le persone da tre anni presenti, un grazie-merci speciale va a *Mariangela* e *Roland.* Supervisors stimolanti, divertenti, indispensabili per la riuscita di questa tesi.
Grazie a tutti i collaboratori e colleghi italo/francesi, importantissimi scientificamente e umanamente: V. Bianciotto, E. Lumini, S. Voyron, F. Barbi, P.Luis, M.P, *Maffa, Albi, Gigi, Marco, G. Russo, Beatris* .
Un grazie particolare a *Fabiano*: ottimo collega, ma soprattutto grande amico in grado, sempre, di farmi stare "*trannquillaaaa*".
Infine, grazie alla mia *famiglia* e agli *amici*, lontani ma sempre vicini e a A. G. per l'estrema pazienza.